

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 03:36:13 : Search time 2066.03 Seconds
(without alignments)
1024.800 Million cell updates/sec

Title: US-08-978-456-1

Perfect score: 1029
Sequence: 1 ATGGATTATGGATTCAACT.....TAAGTTACGAAGAAAGACTGA 1029

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10041192 seqs, 1028798125 residues

Total number of hits satisfying chosen parameters: 20082384

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA.*
1: /cgn2_6/ptodata/2/pna/US082B_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US080_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US081A_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US081B_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US081C_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US081D_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US082A_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US082B_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US082C_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US083A_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US083B_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US084A_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US084B_COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US084C_COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US085A_COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US087A_COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US087B_COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US087C_COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US088A_COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US088B_COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US088C_COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US088D_COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US089A_COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US089B_COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US089C_COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US089D_COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US089E_COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US089F_COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US089G_COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US089H_COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US089I_COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US090A_COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US090B_COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US090C_COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US090D_COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US090E_COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US091A_COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US091B_COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US091C_COMB.seq.*

44: /cgn2_6/ptodata/2/pna/US092B_COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US092C_COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US092D_COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US092E_COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US093A_COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US093B_COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US093C_COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US093D_COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US093E_COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US094A_COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US094B_COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US094C_COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US094D_COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US095_COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US095A_COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US095B_COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US095C_COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US095D_COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US095E_COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US095F_COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US095G_COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US095H_COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US095I_COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US095J_COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US095K_COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US095L_COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US095M_COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US095N_COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US095O_COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US095P_COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US095Q_COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US095R_COMB.seq.*
76: /cgn2_6/ptodata/2/pna/US095S_COMB.seq.*
77: /cgn2_6/ptodata/2/pna/US095T_COMB.seq.*
78: /cgn2_6/ptodata/2/pna/US095U_COMB.seq.*
79: /cgn2_6/ptodata/2/pna/US095V_COMB.seq.*
80: /cgn2_6/ptodata/2/pna/US095W_COMB.seq.*
81: /cgn2_6/ptodata/2/pna/US095X_COMB.seq.*
82: /cgn2_6/ptodata/2/pna/US095Y_COMB.seq.*
83: /cgn2_6/ptodata/2/pna/US095Z_COMB.seq.*
84: /cgn2_6/ptodata/2/pna/US096A_COMB.seq.*
85: /cgn2_6/ptodata/2/pna/US096B_COMB.seq.*
86: /cgn2_6/ptodata/2/pna/US096C_COMB.seq.*
87: /cgn2_6/ptodata/2/pna/US096D_COMB.seq.*
88: /cgn2_6/ptodata/2/pna/US096E_COMB.seq.*
89: /cgn2_6/ptodata/2/pna/US096F_COMB.seq.*
90: /cgn2_6/ptodata/2/pna/US096G_COMB.seq.*
91: /cgn2_6/ptodata/2/pna/US096H_COMB.seq.*
92: /cgn2_6/ptodata/2/pna/US096I_COMB.seq.*
93: /cgn2_6/ptodata/2/pna/US096J_COMB.seq.*
94: /cgn2_6/ptodata/2/pna/US096K_COMB.seq.*
95: /cgn2_6/ptodata/2/pna/US096L_COMB.seq.*
96: /cgn2_6/ptodata/2/pna/US096M_COMB.seq.*
97: /cgn2_6/ptodata/2/pna/US096N_COMB.seq.*
98: /cgn2_6/ptodata/2/pna/US096O_COMB.seq.*
99: /cgn2_6/ptodata/2/pna/US096P_COMB.seq.*
100: /cgn2_6/ptodata/2/pna/US096Q_COMB.seq.*
101: /cgn2_6/ptodata/2/pna/US096R_COMB.seq.*
102: /cgn2_6/ptodata/2/pna/US096S_COMB.seq.*
103: /cgn2_6/ptodata/2/pna/US096T_COMB.seq.*
104: /cgn2_6/ptodata/2/pna/US096U_COMB.seq.*
105: /cgn2_6/ptodata/2/pna/US096V_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1029	100.0	1029	1	PCR-US98-24857-1 Sequence 1, Appl1

1 ATGATATATGCGATTCACCTTGCATAATATGTCACAGGTCAAAGCGTGTATCCACC 60
1 atgattatattgcattcaacttcgaatatagatacgaagccaacgaggttaaccaccc 60
61 GTTGCGCGTGTGTAGTTAATGAAGTAGAGTTGTTGTTATGTTGCACACTTGGAAAA 120
61 gttagcgctgtgttagttaatgaagtagagttgttgcattgtggtgcacacttggaaaa 120
121 GGTGCAAGCATGCGGAGGTTCAAGCACTTGATATGGCAACAAATGCTAAGTGG 180
121 ggtgcaagcatgcgaggttcaagcacttgatggtgcacaaatgctgaagtgg 180
181 ACGATTTATATACGTTAGAGCCATGTAAGTATGTTGTTGTTCAACACACCGTGTAAAC 240
181 acgatttatatacgttagagccatgtagctattgtgtcaacacccggttgaac 240
241 AAAATTATGATTTAGATAGCAAAAGTAGTATACGCAACAAAGCAATTGCTTAGAC 300
241 aaattatattgatttagaagtagcaaaagtagtatacgcacaaaagcaattcgttagac 300
301 ACACATGATGATGACACGTTACGGGCTCAGGTTATGAGGTTGAATGCTTATGATGAA 360
301 acacatgattgatacgttacgagctcaggttatgaaagttagaagtcgttagatgaa 360
421 ACAGTGAAGATGATCTGCAAGTTAGATGCTAAACAGCGAATGATATATGACAAAGTCAA 480
421 acagtgaagatgattctgcaagtttagatgcttaaacagcgaaatgattatgacaaagtc 480
481 TGGATTACTAACAAAGAGTTTAAACAGATGCTATTAATTAAAGATGACATGACACGCA 540
481 tggattactaacaaagagttttaaagagttctataaagttgaacatgacacgacgca 540
541 GGTGTAACGGAAGAGCGTACAGTTGATAGATGATCCACATATATCTACAGCTATTCAA 600
541 ggtgtaacgggaagagcgtaacagtttaataagatgacccaatactacacagattcaa 600
601 GATGAAAAAACCCTATATAAAGTAAATATTTGCTAAGTCTGGAATATTCATTTAATCAG 660
601 gatgaaaaaaccctataaagtaaatattgtctaaagtcgggaatatctatcttaacag 660
721 AGCATCAAAACATATTTGAATTTATTTCTGAGTCTTGATTTAACAACAATCTT 780
721 agcatcaaaacatatttgaattttatcttgaagtccttgatcttaacaacaattctt 780
781 CACAATTTATATAAAGAGAGAGTGAAGTCTGAGTGAAGGAGTCAACCACTACT 840
781 cacaatttatataaagagaggttgaaacttgcctagtcgagggctccacactact 840
841 TCAGATTTCTCATCTATTATATAGTGAATTTATTTCTATATTAAGCCCGAAATTAAT 900
841 tcagatttctcatctatttatatagatgaattatctctatattatccccgaatlaatt 900
901 GCGGATTCGGAATTTATCAATTTTATCAACAATGATGATGATGATGATGATGATGATG 960
901 gcggatctcggaatttatcaattttatcaacaatgattgattgagtlaccagatcg 960
961 AACCAATTTGAATTTGTTTATCCGATTTATTAATCAAAATGTTAAATTAATTAATCA 1020
961 aaccatttgaatttgttcatccgagttatcaaatcaaaaagttaacttaacttaaga 1020
1021 AAGAAGTGA 1029
1021 aagaagtga 1029

RESULT 4
PCT-US98-25068-7
; Sequence 7, Application PC/TUS9825068
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham Corporation
; TITLE OF INVENTION: riba
; FILE REFERENCE: P50444-07
; CURRENT APPLICATION NUMBER: PCT/US98/25068
; EARLIER FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO: 7
; LENGTH: 3336
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
PCT-US98-25068-7

Query Match 100.0%; Score 1029; DB 1; Length 3336;
Best Local Similarity 100.0%; Pred. No. 6e-22e;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGATATATGCGATTCACCTTGCATAATATGTCACAGGTCAAAGCGTGTATCCACC 60
1 atgattatattgcattcaacttcgaatatagatacgaagccaacgaggttaaccaccc 60
61 GTTGCGCGTGTGTAGTTAATGAAGTAGAGTTGTTGTTATGTTGCACACTTGGAAAA 120
61 gttagcgctgtgttagttaatgaagtagagttgttgcattgtggtgcacacttggaaaa 120
121 GGTGCAAGCATGCGGAGGTTCAAGCACTTGATATGGCAACAAATGCTAAGTGG 180
121 ggtgcaagcatgcgaggttcaagcacttgatggtgcacaaatgctgaagtgg 180
181 ACGATTTATATACGTTAGAGCCATGTAAGTATGTTGTTGTTCAACACACCGTGTAAAC 240
181 acgatttatatacgttagagccatgtagctattgtgtcaacacccggttgaac 240
241 AAAATTATGATTTAGATAGCAAAAGTAGTATACGCAACAAAGCAATTGCTTAGAC 300
241 aaattatattgatttagaagtagcaaaagtagtatacgcacaaaagcaattcgttagac 300
301 ACACATGATGATGACACGTTACGGGCTCAGGTTATGAGGTTGAATGCTTATGATGAA 360
301 acacatgattgatacgttacgagctcaggttatgaaagttagaagtcgttagatgaa 360
421 ACAGTGAAGATGATCTGCAAGTTAGATGCTAAACAGCGAATGATATATGACAAAGTCAA 480
421 acagtgaagatgattctgcaagtttagatgcttaaacagcgaaatgattatgacaaagtc 480
481 TGGATTACTAACAAAGAGGTTTAAACAGATGCTATTAATTAAAGATGACATGACACG 540
481 tggattactaacaaagagttttaaagagttctataaagttgaacatgacacgacgca 540
541 GGTGTAACGGAAGAGCGTACAGTTGATAGATGATCCACATATATCTACAGCTATTCAA 600
541 ggtgtaacgggaagagcgtaacagtttaataagatgacccaatactacacagattcaa 600
601 GATGAAAAAACCCTATATAAAGTAAATATTTGCTAAGTCTGGAATATTCATTTAATCAG 660
601 gatgaaaaaaccctataaagtaaatattgtctaaagtcgggaatatctatcttaacag 660
661 CAAATTTATCAAGATGAATCAACACCAATTTGATATATGATGAAATCCAAATTTAACA 720
661 caaatttcaagatgaatcaacacccaatttggatatatactatcgaanaatccaatttaaca 720


```
QY 1 ATGATTATGCGATTCACCTTGCATATATGTGACAGGTCACAAAGAGTGTATATCCACC 60
Db 1 atgattatgcgattcacctaacttgcataatgtgtacaaagttcaaacgggttcaatccacc 60
QY 61 GTTGCGCGTGTGTAGTTAATGAAGTAGATTTGTGTATTTGTGCACACTTGAGAAA 120
Db 61 gttgcgcggtgtgttagttaatgaagtagatgtgtgtatgtgtgcacacttgagaaa 120
QY 121 GGTGACAGCATGCGGAGGCTTCAACACTTGTATATGACCAACAACAAATGCTGAGAGTCG 180
Db 121 ggtgacagcatgcgagagcttcaacacttgalatgacacacaaatgtcgaagtgcg 180
QY 181 AGCATTTATATACGTGTAGAGCCATGTAGTCATTTTGTTCACACACCCCTGTGTAC 240
Db 181 agcatttatatacgtttagagccatgtatgtcatcttgglttcaaccaccctgtgttac 240
QY 241 AAAATTATGTATTGTAGATAGCAAAAGTAGTATACGACAAAAGACAAATTCGTTAGAC 300
Db 241 aaaattatgtattgttagatagcaaaagtagatatacgcaacaaagacaattcgttagac 300
QY 301 ACACATGCTGATGAGACGTTACGGGCTCACGCTATTTGAGGTTGATGCGTTGATGATGAA 360
Db 301 acacatgctgatgagacgtttacgggctcacgctatgttgaagtgctgagatgaa 360
QY 361 CGGCATCACAAATTTATACCAAGACTTTTAAAGCAAAAGCAAAAGCAACTGCGACAATTT 420
Db 361 cggcatcacaaatttataccagactttttaaagcaaaaagcaaaagcaactgcacaaatt 420
QY 421 ACAGTGAAGATATCTGCAAGTTTATGATGTAACAAGCGCAATGATATGACAAAAGTCAA 480
Db 421 acagtgaagatctctgcaagttttagatgtataacaagcgaaatgataatgcaaaagtc 480
QY 481 TGGATTACTAACAAGAGGTTAAACAAGATGCTTAAGTATAGCATTCGACACACCA 540
Db 481 tggattactaacaagaggtttaaacaagatgcttaagtaagcaccgcgacacacca 540
QY 541 GTGTTAACTGAGAGACGTACAGTTGAATTAGTATCCAAATTAATCTACAGTATTCAA 600
Db 541 gtgttaactgagagacgtacagttgaattagatgatacgcacaatactacacgtattcaa 600
QY 601 GATGAAAAAACCTTAATAAAGTAATTTGTCTAAAGTGTGGAAATATTCATTTAATCAG 660
Db 601 gatgaaaaaaccttaataaagtaattgtctaaagtgtggaaatattcatcttaacag 660
QY 661 CAATTTTCAAGATGATCAACACCAATTTGATATATGCAAAATCCAAATTTAACA 720
Db 661 caattttcaagatgatacaccacaaatltgatatatactgaaaaatccaaatltaca 720
QY 721 AGCAATCAACACATATGAAATTTTACTTGAAGTCTGTGATTTTAAACAATCTTT 780
Db 721 agcaatcaaacacatattgaaatttctacttgaagctcttgyatlttaacaaacttct 780
QY 781 CACAATTTATATAAAGAGAGTGGAACTTTGCTAGTCGAGGCGTCCAACTACT 840
Db 781 cacaatttatataaagagagtggaacttgcctagtcgagcgccaaacccactact 840
QY 841 TCAGAAATCTCCATCTATATATATAGATGAATTTATCTCTATTAAGCCCCCAATTAAT 900
Db 841 tcagaatctccatctatataatataatgaatltatctctatagcccccgaatlaatt 900
QY 901 GCGGATCTGGAATTTATCAATTTTATCAAAACAAATGATGATGATGATGATGATGATG 960
Db 901 gcggatctggaatttatcaattttatacaaaatgagtgtgatatgagataccagatcg 960
QY 961 AACCAATTTGAATTTGTTCAATTCGAGTTATTAATCAAAATGTTAAATTAATTAAC 1020
Db 961 aaccaatttgaatttgttcattccgagttatataatcaaaatgttaactttagca 1020
QY 1021 AAGAAGTGA 1029
Db 1021 aagaagtga 1029
```

```
RESULT 7
US-08-977-553-5
: Sequence 5: Application US/08977553
: GENERAL INFORMATION:
: APPLICANT: Palmer, Leslie M.
: TITLE OF INVENTION: Novel ribh
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechert Price & Rhoads
: STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
: CITY: Philadelphia
: STATE: PA
: COUNTRY: US
: ZIP: 19103
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Dickinson, Todd O
: REGISTRATION NUMBER: 28,354
: REFERENCE/DOCKET NUMBER: GM10123
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 215-994-2252
: TELEFAX: 215-994-2222
: TELEX:
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3336 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: US-08-977-553-5

Query Match 100.0%; Score 1029; DB 32; Length 3336;
Best Local Similarity 100.0%; Pred. No. 66-224;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATTATGCGATTCACCTTGCATATATGTGACAGGTCACAAAGAGTGTATATCCACC 60
Db 1 atgattatgcgattcacctaacttgcataatgtgtacaaagttcaaacgggttcaatccacc 60
QY 61 GTTGCGCGTGTGTAGTTAATGAAGTAGATTTGTGTATTTGTGCACACTTGAGAAA 120
Db 61 gttgcgcggtgtgttagttaatgaagtagatgtgtgtatgtgtgcacacttgagaaa 120
QY 121 GGTGACAGCATGCGGAGGCTTCAACACTTGTATATGACCAACAACAAATGCTGAGAGTCG 180
Db 121 ggtgacagcatgcgagagcttcaacacttgalatgacacacaaatgtcgaagtgcg 180
QY 181 AGCATTTATATACGTGTAGAGCCATGTAGTCATTTTGTTCACACACCCCTGTGTAC 240
Db 181 agcatttatatacgtttagagccatgtatgtcatcttgglttcaaccaccctgtgttac 240
QY 241 AAAATTATGTATTGTAGATAGCAAAAGTAGTATACGACAAAAGACAAATTCGTTAGAC 300
Db 241 aaaattatgtattgttagatagcaaaagtagatatacgcaacaaagacaattcgttagac 300
QY 301 ACACATGCTGATGAGACGTTACGGGCTCACGCTATTTGAGGTTGATGCGTTGATGATGAA 360
Db 301 acacatgctgatgagacgtttacgggctcacgctatgttgaagtgctgagatgaa 360
QY 361 CGGCATCACAAATTTATACCAAGACTTTTAAAGCAAAAGCAAAAGCAACTGCGACAATTT 420
Db 361 cggcatcacaaatttataccagactttttaaagcaaaaagcaaaagcaactgcacaaatt
```

Db 361 CGGCGATACAAATATATACCAAGACTTTTAAAGCAAAAGCAACCACTGCCACAATT 420
Qy 421 ACAGTGAATATCTGCAATTTAGATGTAAACACGATGATATATGACAAAGTCAA 480
Db 421 ACAGTGAATATCTGCAATTTAGATGTAAACACGATGATATATGACAAAGTCAA 480
Qy 481 TGGATTACTAACAAAGGTTAAACAAGATGTATAGTTAAACATCAGACGACGA 540
Db 481 TGGATTACTAACAAAGGTTAAACAAGATGTATAGTTAAACATCAGACGACGA 540
Qy 541 GTGTTAACTGGAAGACGTACAGTTGAATTAAGATGATCCACAATATCTACAGTATCAA 600
Db 541 GTGTTAACTGGAAGACGTACAGTTGAATTAAGATGATCCACAATATCTACAGTATCAA 600
Qy 601 GATGGAAGAAACCTATATAAGTATATGCTAGCTGGGAATATCTCAATTTAATCAG 660
Db 601 GATGGAAGAAACCTATATAAGTATATGCTAGCTGGGAATATCTCAATTTAATCAG 660
Qy 661 CAATTTATCAGATGATCAACACCAATTTGATATATCTGAAATCCAAATTTTACA 720
Db 661 CAATTTATCAGATGATCAACACCAATTTGATATATCTGAAATCCAAATTTTACA 720
Qy 721 AGCAATCAACACATATTTGAATTTATTTACTGTAAGTCTTGATTTAACACAAATCTT 780
Db 721 AGCAATCAACACATATTTGAATTTATTTACTGTAAGTCTTGATTTAACACAAATCTT 780
Qy 781 CACATTTATATAAAGAGAGAGTGGAACTTGTCTAGTCCAGCAGGTCACACCACTACT 840
Db 781 CACATTTATATAAAGAGAGAGTGGAACTTGTCTAGTCCAGCAGGTCACACCACTACT 840
Qy 841 TCAGATTTCCATCTATTTATATAGATGATTTATTTCTATTTATGCCCCGAATTTAT 900
Db 841 TCAGATTTCCATCTATTTATATAGATGATTTATTTCTATTTATGCCCCGAATTTAT 900
Qy 901 GCGGCGATCTGGAATTTATCAATTTTATCAACAACATGATGATGAGATCCAGATGCG 960
Db 901 GCGGCGATCTGGAATTTATCAATTTTATCAACAACATGATGATGAGATCCAGATGCG 960
Qy 961 AACCAATTTGAATTTGTCATTCGAGATTTAATCAAAATGTAATTAATTAATTAATTCGA 1020
Db 961 AACCAATTTGAATTTGTCATTCGAGATTTAATCAAAATGTAATTAATTAATTAATTCGA 1020
Qy 1021 AAGAGTGA 1029
Db 1021 AAGAGTGA 1029

RESULT 8
US-08-977-866-7
: Sequence 7. Application US/08977866
: GENERAL INFORMATION:
: APPLICANT: Palmer, Leslie M.
: TITLE OF INVENTION: Novel ribB
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechert Price & Rhoads
: STREET: 4000 Bell Atlantic Tower, 1717 Arch St
: CITY: Philadelphia
: STATE: PA
: COUNTRY: US
: ZIP: 19103
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/977,866
: FILING DATE:
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:

: ATTORNEY/AGENT INFORMATION:
: NAME: Dickinson, Todd O
: REGISTRATION NUMBER: 28,354
: REFERENCE/DOCKET NUMBER: P50444-8
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 215-994-2252
: TELEFAX: 215-994-2222
: TELEX:
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3336 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: US-08-977-866-7

Query Match 100.0%; Score 1029; DB 32; Length 3336;
Best Local Similarity 100.0%; Pred. No. 6e-224;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGATTTATGCGATTCACCTGCAATTTATGTTCAAGTTCACAGCTGTAATCCACCC 60
Db 1 ATGATTTATGCGATTCACCTGCAATTTATGTTCAAGTTCACAGCTGTAATCCACCC 60
Qy 61 GTTGGCGCTGTGTAGTTAATGAAGTAGGATGTTGTTATGTCACACTGAGAGAAA 120
Db 61 GTTGGCGCTGTGTAGTTAATGAAGTAGGATGTTGTTATGTCACACTGAGAGAAA 120
Qy 121 GGTGACAAAGCATCGGAGGTTCAAGCATTTGATATGTCACAAACAAATGCTGAGGTGG 180
Db 121 GGTGACAAAGCATCGGAGGTTCAAGCATTTGATATGTCACAAACAAATGCTGAGGTGG 180
Qy 181 AGCATTTATTTACGTTAGAGCCATGATGATTTTGGTTCAACACCACCTGTGTTAAC 240
Db 181 AGCATTTATTTACGTTAGAGCCATGATGATTTTGGTTCAACACCACCTGTGTTAAC 240
Qy 241 AAAATTTATGATTTGTAAGTATGCAAAAGTATGATACGCAACAAAGCAATTCGTTAGAC 300
Db 241 AAAATTTATGATTTGTAAGTATGCAAAAGTATGATACGCAACAAAGCAATTCGTTAGAC 300
Qy 301 ACACATGTTGATGAGACGTTACGGGCTCAACGGTATTTGAGGTTGGAATGCGTTGATGAA 360
Db 301 ACACATGTTGATGAGACGTTACGGGCTCAACGGTATTTGAGGTTGGAATGCGTTGATGAA 360
Qy 361 CGGCGATCACAAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCAACTGCCACAAATT 420
Db 361 CGGCGATCACAAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCAACTGCCACAAATT 420
Qy 421 ACAGTGAAGTATCTGCAAGTTAGATGTAACAAAGCGAATGATATAGACAAAGTCAA 480
Db 421 ACAGTGAAGTATCTGCAAGTTAGATGTAACAAAGCGAATGATATAGACAAAGTCAA 480
Qy 481 TGGATTACTAACAAAGGTTAAACAAGATGTCTATAGTTAAAGATGATGACGACGACA 540
Db 481 TGGATTACTAACAAAGGTTAAACAAGATGTCTATAGTTAAAGATGATGACGACGACA 540
Qy 541 GTGTTAACTGGAAGACGTACAGTTGAATTAAGATGATCCACAATATCTACAGTATCAA 600
Db 541 GTGTTAACTGGAAGACGTACAGTTGAATTAAGATGATCCACAATATCTACAGTATCAA 600
Qy 601 GATGGAAGAAACCTATATAAGTATATGCTAGCTGGGAATATCTCAATTTAATCAG 660
Db 601 GATGGAAGAAACCTATATAAGTATATGCTAGCTGGGAATATCTCAATTTAATCAG 660
Qy 661 CAATTTATCAGATGATCAACACCAATTTGATATATCTGAAATCCAAATTTTACA 720
Db 661 CAATTTATCAGATGATCAACACCAATTTGATATATCTGAAATCCAAATTTTACA 720
Qy 721 AGCAATCAACACATATTTGAATTTATTTACTGTAAGTCTTGATTTAACACAAATCTT 780
Db 721 AGCAATCAACACATATTTGAATTTATTTACTGTAAGTCTTGATTTAACACAAATCTT 780

OY 1021 AAGACTGA 1029
Db 1021 AAGACTGA 1029

RESULT 10
US-09-227-806-7
Sequence 7, Application US/09227806
GENERAL INFORMATION:
APPLICANT: Palmer, Leslie M.
APPLICANT: Fedon, Jason C.
APPLICANT: Warren, Richard L.
APPLICANT: Trainl, Christopher M.
APPLICANT: Wang, Min
APPLICANT: Jaworski, Deborah D.
APPLICANT: Mooney, Jeffrey
APPLICANT: Debouck, Christine
APPLICANT: Zhong, Yixi
APPLICANT: Black, Michael
TITLE OF INVENTION: r1bA
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dechert, Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2793
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/227,806
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/977,554
FILING DATE:
APPLICATION NUMBER: PCT/US97/02318
FILING DATE: 19-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, O. Todd
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50444-07
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215/994-2252
TELEFAX: 215/994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3336 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-227-806-7

Query Match 100.0%; Score 1029; DB 43; Length 3336;
Best Local Similarity 100.0%; Pred. No. 6e-224;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGATATGCGATTCACCTTGCACAAATATGTAACAAGTCAACAGGTGTTATCCACC 60
Db 1 ATGATATGCGATTCACCTTGCACAAATATGTAACAAGTCAACAGGTGTTATCCACC 60

OY 61 GTTGGCGCTGTGTACTATATGAAGTAGATGTGTGTTATGTCACACTTGAGAAAA 120
Db 61 GTTGGCGCTGTGTACTATATGAAGTAGATGTGTGTTATGTCACACTTGAGAAAA 120

OY 121 GGTGACAACATCGGAGGTTCAAGCACTTGATATGCAACAACAAATGCTGAAGGTGG 180

Db 121 GGTGACAACATCGGAGGTTCAAGCACTTGATATGCAACAACAAATGCTGAAGGTGG 180

OY 181 ACGATTATATTACGTAGAGCCATGTAGTCAATTTGGTTCAACACCACTGTGTTAAC 240
Db 181 ACGATTATATTACGTAGAGCCATGTAGTCAATTTGGTTCAACACCACTGTGTTAAC 240

OY 241 AAAATTATGATGTAAAGTACCAAAAGTAGATGCGCAACAAAGACAATTCGTTAAC 300
Db 241 AAAATTATGATGTAAAGTACCAAAAGTAGATGCGCAACAAAGACAATTCGTTAAC 300

OY 301 ACACATGCTGATAGACGCTTACGCGCTCACGGTATGAGTGAATGCGTTGATGATGA 360
Db 301 ACACATGCTGATAGACGCTTACGCGCTCACGGTATGAGTGAATGCGTTGATGATGA 360

OY 361 CGGCGATCACAATTTATACCAAGACTTTTAAAGCAAAAGCAACCACTGCCACAAAT 420
Db 361 CGGCGATCACAATTTATACCAAGACTTTTAAAGCAAAAGCAACCACTGCCACAAAT 420

OY 421 ACAGTGAAGTATCTGCAAGTTAGATGTTAAACAAGGGAATGATATGACAAAGTCAA 480
Db 421 ACAGTGAAGTATCTGCAAGTTAGATGTTAAACAAGGGAATGATATGACAAAGTCAA 480

OY 481 TCGATTACTAACAAAGAGGTTAAACAAGATGCTATTAAGTTAGACATCGACAGCACA 540
Db 481 TCGATTACTAACAAAGAGGTTAAACAAGATGCTATTAAGTTAGACATCGACAGCACA 540

OY 541 GTGTTAACTGAAGAAGTACAGTGAATAGATGATCCACAATATACTACAGTATTCAA 600
Db 541 GTGTTAACTGAAGAAGTACAGTGAATAGATGATCCACAATATACTACAGTATTCAA 600

OY 601 GATGGAAGAAACCTATTAAGTATATGCTAGCTGGGAATATCATTTATATAG 660
Db 601 GATGGAAGAAACCTATTAAGTATATGCTAGCTGGGAATATCATTTATATAG 660

OY 661 CAAATTTATCAAGATGAATCAACACCAATTGATATAGTGAATCAAAATTTAAACA 720
Db 661 CAAATTTATCAAGATGAATCAACACCAATTGATATAGTGAATCAAAATTTAAACA 720

OY 721 AGCAATCAACACATATTTGAATTTATTTACTGGAAGCTTGTGATTTAAACAATTTCT 780
Db 721 AGCAATCAACACATATTTGAATTTATTTACTGGAAGCTTGTGATTTAAACAATTTCT 780

OY 781 CACAATTTATATAAAGAGGAGTGTGAACCTTGTAGTGAAGGCGCAACCACTACT 840
Db 781 CACAATTTATATAAAGAGGAGTGTGAACCTTGTAGTGAAGGCGCAACCACTACT 840

OY 841 TCAGAAATTCCTCATCTATTTATATAGATGAATTTATCTCTATTATGCCCCGAATTAAT 900
Db 841 TCAGAAATTCCTCATCTATTTATATAGATGAATTTATCTCTATTATGCCCCGAATTAAT 900

OY 901 GCGGATCTGGAATTTATCAATTTTATCAACAATGATGTGATGACATACAGATCG 960
Db 901 GCGGATCTGGAATTTATCAATTTTATCAACAATGATGTGATGACATACAGATCG 960

OY 961 AACCAATTTGAATGTTCATTCGAGTTATTAATCAAAATGTTAAATTAACCTTACGA 1020
Db 961 AACCAATTTGAATGTTCATTCGAGTTATTAATCAAAATGTTAAATTAACCTTACGA 1020

OY 1021 AAGACTGA 1029
Db 1021 AAGACTGA 1029

RESULT 11
US-09-492-479-7
Sequence 7, Application US/09492479
GENERAL INFORMATION:
APPLICANT: Palmer, Leslie M.
TITLE OF INVENTION: Novel r1bB
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dechert Price & Rhoads

```

STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/492.479
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/977,866
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Todd Q
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50444-8
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2252
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3336 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-492-479-7

```

```

Query Match      100.0%; Score 1029; DB 54; Length 3336;
Best Local Similarity 100.0%; Pred. No. 6e-224;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 ATGGATTATGGATTCAATTCGCAAAATATGGTACAGGTCAACAGGTGTAATCCACC 60
DB 1 ATGGATTATGGATTCAATTCGCAAAATATGGTACAGGTCAACAGGTGTAATCCACC 60
QY 61 GTTGCGCGCTGTTAGTAAATGAAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 120
DB 61 GTTGCGCGCTGTTAGTAAATGAAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 120
QY 121 GGTACAACAGCATGCGGAGGTTCAAGCACTTATATGSCAACAAATGCTGAAGTGG 180
DB 121 GGTACAACAGCATGCGGAGGTTCAAGCACTTATATGSCAACAAATGCTGAAGTGG 180
QY 181 ACGATTATATAGTTAGTACAGCCATGTGTCATTTGTTCAACACCCCTGTGTTAAC 240
DB 181 ACGATTATATAGTTAGTACAGCCATGTGTCATTTGTTCAACACCCCTGTGTTAAC 240
QY 241 AAAATTATGATTGTAAGATAGCAAAATAGTATACGCAACAAAGCAATTCGTTAGAC 300
DB 241 AAAATTATGATTGTAAGATAGCAAAATAGTATACGCAACAAAGCAATTCGTTAGAC 300
QY 301 ACACATGCTGATGAGACGTTACGGGCTACGGTATGATGAGTTGAATGCGTGTGATGAC 360
DB 301 ACACATGCTGATGAGACGTTACGGGCTACGGTATGATGAGTTGAATGCGTGTGATGAC 360
QY 361 CGGCAATCACAATATACCAAGACCTTTTAAAGCAAAAGCAACCTGCCCAATTT 420
DB 361 CGGCAATCACAATATACCAAGACCTTTTAAAGCAAAAGCAACCTGCCCAATTT 420
QY 421 ACGATGAAGTATCTGCAAGTTTAGATGTAACAGCGAATGATTAATGACCAAGTCAA 480
DB 421 ACGATGAAGTATCTGCAAGTTTAGATGTAACAGCGAATGATTAATGACCAAGTCAA 480
QY 481 TGGATTACTAACAAGAGGTTAAACAGATGCTATTAAGATTAAAGCATGACACGACGA 540
DB 481 TGGATTACTAACAAGAGGTTAAACAGATGCTATTAAGATTAAAGCATGACACGACGA 540

```

```

QY 541 GTGTTACTGGAAGACGTACAGTTGAATTAGATGATCACAATATACCTATTCAC 600
DB 541 GTGTTACTGGAAGACGTACAGTTGAATTAGATGATCACAATATACCTATTCAC 600
QY 601 GATGAAAAAACCCTATTAAGTAATTTGCTAGTCTGGAAATTTCAATTTTAAC 660
DB 601 GATGAAAAAACCCTATTAAGTAATTTGCTAGTCTGGAAATTTCAATTTTAAC 660
QY 661 CAAATTATATCAAGTATATACACCAATTTGATATATATCTGAAAAATCCAAATTTAC 720
DB 661 CAAATTATATCAAGTATATACACCAATTTGATATATATCTGAAAAATCCAAATTTAC 720
QY 721 ACGATCAACACATATTTGAATTTTACTTGAAGCTGTGATTTAACACAATTTCT 780
DB 721 ACGATCAACACATATTTGAATTTTACTTGAAGCTGTGATTTAACACAATTTCT 780
QY 781 CACAATTTATATAAAGAGAGTTGGAACCTTCTAGTCGAGGACAGTCCAACCTACT 840
DB 781 CACAATTTATATAAAGAGAGTTGGAACCTTCTAGTCGAGGACAGTCCAACCTACT 840
QY 841 TCAGAAATTCCTCATCTTATATAGATGAATTTATTTCTATATATGCCCCGAATTAAT 900
DB 841 TCAGAAATTCCTCATCTTATATAGATGAATTTATTTCTATATATGCCCCGAATTAAT 900
QY 901 GCGGATCTGGAATTTATCAATTTTATCAACAAATGATGTGATGATACAGATGCG 960
DB 901 GCGGATCTGGAATTTATCAATTTTATCAACAAATGATGTGATGATACAGATGCG 960
QY 961 AACCAATTTGAATTTGTTCAATTCGAGTATTAATCAAAATGTAATTAACCTTACGA 1020
DB 961 AACCAATTTGAATTTGTTCAATTCGAGTATTAATCAAAATGTAATTAACCTTACGA 1020
QY 1021 AAGAGTGA 1029
DB 1021 AAGAGTGA 1029

```

```

RESULT 12
US-09-376-633-7

```

```

Sequence 7, Application US/09376633
GENERAL INFORMATION:
APPLICANT: Palmer, Leslie M.
TITLE OF INVENTION: Novel ribg
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/376,633
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/978,456
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Todd Q
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50444-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2252
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 7:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 3336 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-376-633-7

```

```

Query Match      100.0%; Score 1029; DB 103; Length 3336;
Best Local Similarity 100.0%; Pred. No. 6e-224;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 ATGATATATCGATTCACTTGCATATATGATGATCAAGTCAAGGATGTTATCCACC 60
    |||||||
DB 1 ATGATATATCGATTCACTTGCATATATGATGATCAAGTCAAGGATGTTATCCACC 60

QY 61 GTTGGCCCTGTTAGTAAATGAAGTAGAGATTGTTGTTGTCACACTTGAGAAAA 120
    |||||||
DB 61 GTTGGCCCTGTTAGTAAATGAAGTAGAGATTGTTGTTGTTGTCACACTTGAGAAAA 120

QY 121 GGTGACAGAGATGGGAGGTTCAAGCACTTGATGTCACACAAAATGCTGAAGTGGC 180
    |||||||
DB 121 GGTGACAGAGATGGGAGGTTCAAGCACTTGATGTCACACAAAATGCTGAAGTGGC 180

QY 181 ACGATTATATACGTAGAGCCATGATGATCTTGTTCACACACCCTGTGTAAAC 240
    |||||||
DB 181 ACGATTATATACGTAGAGCCATGATGATCTTGTTCACACACCCTGTGTAAAC 240

QY 241 AAAATTATGATTGTAAGATAGCAAAAGTAGATACGCAACAAAAGCAATTCGTAGAC 300
    |||||||
DB 241 AAAATTATGATTGTAAGATAGCAAAAGTAGATACGCAACAAAAGCAATTCGTAGAC 300

QY 301 ACACATGCTATGAGAGTACCGGCTCAGCGTATGAGTGAATGCGTTGATGATGAA 360
    |||||||
DB 301 ACACATGCTATGAGAGTACCGGCTCAGCGTATGAGTGAATGCGTTGATGATGAA 360

QY 361 CGGGCATCACAAATATACCAAGACTTTTAAAGCAAAAGCAAGCAACTGCCAATTT 420
    |||||||
DB 361 CGGGCATCACAAATATACCAAGACTTTTAAAGCAAAAGCAAGCAACTGCCAATTT 420

QY 421 ACAAGTAAGTATCTGCAAGTTAGATGTTAACACAGCAATGATGACAAAGTCAA 480
    |||||||
DB 421 ACAAGTAAGTATCTGCAAGTTAGATGTTAACACAGCAATGATGACAAAGTCAA 480

QY 481 TGGATTCTACAAAGAGGTTAAACAGATGCTATAGTTAAACATCCACAGAGGCA 540
    |||||||
DB 481 TGGATTCTACAAAGAGGTTAAACAGATGCTATAGTTAAACATCCACAGAGGCA 540

QY 541 GTGTTACTGAGAGACGTACAGTGAATAGATGATCCACAATATACAGATTTCAA 600
    |||||||
DB 541 GTGTTACTGAGAGACGTACAGTGAATAGATGATCCACAATATACAGATTTCAA 600

QY 601 GATGAAAAAACCTATAAAAAGTAAATGCTAAGCTGGGAATATTCATTTTAATCAG 660
    |||||||
DB 601 GATGAAAAAACCTATAAAAAGTAAATGCTAAGCTGGGAATATTCATTTTAATCAG 660

QY 661 CAATTTATCAAGTGAATCAACCAATTTGATATATCTGAAAAATCCAAATTTAACA 720
    |||||||
DB 661 CAATTTATCAAGTGAATCAACCAATTTGATATATCTGAAAAATCCAAATTTAACA 720

QY 721 AGCAATCAACACATATGTAATATTTACTGTAAGTCTGTGATTTAACACATTTCT 780
    |||||||
DB 721 AGCAATCAACACATATGTAATATTTACTGTAAGTCTGTGATTTAACACATTTCT 780

QY 781 CACAATTTATATAAAGAGAGTGAACCTTGTAGTCGAGCAGGTCCAAACCACTACT 840
    |||||||
DB 781 CACAATTTATATAAAGAGAGTGAACCTTGTAGTCGAGCAGGTCCAAACCACTACT 840

QY 841 TCAGAAATTCATCTATTTATATGATGAATTTATCTCTATTTATGCCCCGAATTAAT 900
    |||||||
DB 841 TCAGAAATTCATCTATTTATATGATGAATTTATCTCTATTTATGCCCCGAATTAAT 900

QY 901 GGGGATCTGGAATATATCAATTTATCAACAAATGATGATGATACAGATGCG 960
    |||||||

```

```

DB 901 GGGGATCTGGAATATATCAATTTATCAACAAATGATGATGATGATACAGATGCG 960
    |||||||
QY 961 AACCAATTTAAATGTTCAATTCGAGATTTAAATCAAAATGTTAAATTAATTTACGA 1020
    |||||||
DB 961 AACCAATTTAAATGTTCAATTCGAGATTTAAATCAAAATGTTAAATTAATTTACGA 1020

QY 1021 AAGAGTGA 1029
    |||||||
DB 1021 AAGAGTGA 1029

```

RESULT 13

```

US-08-827-356-1376
; Sequence 1376 Application US/08827356
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Robert S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS RELATED
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 5574
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corporation
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033-0530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/827,356
; FILING DATE: 01-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/014,477
; FILING DATE: 01-APR-1996
; APPLICATION NUMBER: 60/016,743
; FILING DATE: 02-MAY-1996
; APPLICATION NUMBER: 60/020,016
; FILING DATE: 14-JUN-1996
; INFORMATION FOR SEQ. ID NO: 1376:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1119 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1...1119
US-08-827-356-1376

```

```

Query Match      97.8%; Score 1006.4; DB 23; Length 1119;
Best Local Similarity 99.1%; Pred. No. 6.2e-219;
Matches 1023; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

```

```

QY 1 ATGATATATCGATTCACTTGCATATATGATGATCAAGTCAAGGATGTTATCCACC 60
    |||||||
DB 87 ATGATATATCGATTCACTTGCATATATGATGATCAAGTCAAGGATGTTATCCACC 146

QY 61 GTTGGCCCTGTTAGTAAATGAAGTAGAGATTGTTGTTGTCACACTTGAGAAAA 120
    |||||||
DB 147 GTTGGCCCTGTTAGTAAATGAAGTAGAGATTGTTGTTGTTGTCACACTTGAGAAAA 206

```



```
Db 208 ggtgacaacatcgagggttcaagcattgatggtcacacaaaatgctgaagtgcg 267
Oy 181 ACGATTATATTCGTTAGAGCCATGAGTCATTTGGTTCACACACCCTGTTTAC 240
Db 268 acgattatattcgttagagccatgtagcatttggttcaacacacacctgtgtaac 327
Oy 241 AAAATTATGATTGTAAAGTAGCAAAAAGTAGTACGCAACAAAGCAATTCGTTAGAC 300
Db 328 aaattatttgattgtagaagtagcaaaaagttagatagcgaacaaagcaattcgtagac 387
Oy 301 ACACATGCTGATGACAGCTTACGGGCTCAGCGTATTGAGGTTGAGCGTTGATGAA 360
Db 388 acacatggtgatgagacgttaccggtccacgatttagagttgacgttgatgtaa 447
Oy 361 CGGCGATTCACAAATTATACAGAGCTTTTAAAGCAAAAGCAACGACGCAAAATT 420
Db 448 cgggcatcacattatataccaagaccttttaagcaaaagcaagcaacgcaaatc 507
Oy 421 ACAGTGAAGTATCTGCAAGTTAGATGTTAAACAAGCGATGTAATGAGCAAACTCAA 480
Db 508 acagtgaaagtatctgcaagcttagatggttaaacgaagcgaatgataatgcaaaagtc 567
Oy 481 TGCAATTACTACAAAGAGTTAAACAAGATGCTATATAGTTAAGACATGCACAGACGCA 540
Db 568 tggattactacaagaaggtttaaacaagaatgctatagtttaagacacgcagcagca 627
Oy 541 GTTTAACTGAGAGCTAGAGTTGAATTGATGATCCACAATATCTACAGTATTCAA 600
Db 628 gtttaactggaagacgttaagttgaatgatacacaatactacacgtatcaaa 687
Oy 601 GATGGAAGAAACCCCTATATAAGTAATTTGCTAAGTCTGGGAATATTCATTTAATCAG 660
Db 688 gatggaagaaacacctataaagaataatgtctcaagctcgtggaatatacttttaacag 747
Oy 661 CAATTTATCAGATGATGATACACCAATTTGGATATATCTGAAAAATCCAAATTTACA 720
Db 748 caatttatcaagatgaatcaacaacaaatttgatatactgaaaaatccaaatttaca 807
Oy 721 AGCAATCAACACATATTTGAATTTATTTACTGAGCTTGATTTAAACAATCTT 780
Db 808 agaatcaaacacatattgaaattacttgaagctcgtgtaatttaaaaaaatctt 867
Oy 781 CACAAATTTATATAAGAGAGAGTGGAACTTGTAGTCGAGCAGGTCCACACTACT 840
Db 868 caaatattatataaagagagtgtagacttgtagcagcaggtccaaacacact 927
Oy 841 TCAGAAATTCCTCATCT--ATTATATGATGAATTTATTTCTATTTATGCCCCGAAATTA 897
Db 928 tcgaatttctccaatctaatatataatagaaatttattctctattatgcccgaattta 987
Oy 898 ATGGCGGATCTGGAATTTATCAATTTTATCAACAAATGATGTGATGATACCAAT 957
Db 988 attgagcgagatctggaatttcaatttatacaacaagaatgtagatgataccagat 1047
Oy 958 GCGAACCAATTTGAATTTGTTTCGAGATTATTAATCAAAATGTTAATTAATCTTA 1017
Db 1048 ggcgaaccaattgtaaatgttccatccgagttatataatcaaaaatgtaataactta 1107
Oy 1018 CGAAGAGAGTGA 1029
Db 1108 cgaagaagagtgta 1119
```

```
RESULT 15
US-09-611-529-7446/c
; Sequence 7446, Application US/09611529
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Roberta S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
; FILE REFERENCE: 1034/IC963US1
```

```
; CURRENT APPLICATION NUMBER: US/09/611,529
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/417,811
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/353,718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,334
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,221
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,137
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,082
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,081
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,079
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/035,913
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/035,744
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 08/827,356
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: US 08/831,156
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: US 60/014,477
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: US 60/016,743
; PRIOR FILING DATE: 1996-05-02
; PRIOR APPLICATION NUMBER: US 60/020,016
; PRIOR FILING DATE: 1996-06-14
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 7446
; LENGTH: 74105
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-611-529-7446

Query Match          97.8%  Score 1006.4:  DB 57:  Length 74105:
Best Local Similarity 99.1%:  Pred. No. 2e-218:
Matches 1023:  Conservative 0:  Mismatches 6:  Indels 3:  Gaps 1:
```

```
Oy 1 ATGATTATGCGATTCACTTGCATAATATGTACAGAGTCAAAACAGGTGTTATCCACC 60
Db 24245 ATGATTATGCGATTCACTTGCATAATATGTACAGAGTCAAAACAGGTGTTATCCACC 24186
Oy 61 GTTGGCGCTGTTGTTACTTAATGAGGTAGAGATTGTTGTTATGTTGTCACACTTGAGAAAA 120
Db 24185 GTTGGCGCTGTTGTTACTTAATGAGGTAGAGATTGTTGTTATGTTGTCACACTTGAGAAAA 24126
Oy 121 GGTGACAAGATCGGAGGTTCACAGCACTTGATGATGCAACCAACAAATGCTGAAGTGGC 180
Db 24125 GGTGACAAGATCGGAGGTTCACAGCACTTGATGATGCAACCAACAAATGCTGAAGTGGC 24066
Oy 181 ACGATTATATTCGTTAGAGCCATGAGTCATTTGGTTCACACACCCTGTGTTAAC 240
```

```
|||||
Db 24065 ACGATTTATATAGCTTAGACCAATGATGTTGGTTCACACACCCTGTGTAC 24006
QY 241 AAAATTTATGTTGAATAGCAAAAGTAGATACGCAACAAAAGCAATTCGTTAGAC 300
Db 24005 AAAATTTATGTTGAATAGCAAAAGTAGATAGCAACAAAAGCAATTCGTTAGAC 23946
QY 301 ACACATGGTGTAGAGACGCTTACGGGCTCAGGTAATGAGTTGAATGCGTTGATGAA 360
Db 23945 ACACATGGTGTAGAGACGCTTACGGGCTCAGGTAATGAGTTGAATGCGTTGATGAA 23886
QY 361 CGGCATCACAATTATACCAAGACTTTTAAAGCAAAAGCAAACTGCCACAAT 420
Db 23885 CGGCATCACAATTATACCAAGACTTTTAAAGCAAAAGCAAACTGCCACAAT 23826
QY 421 ACAGTGAAGTATCTGCAAGTTTGAATGTAACAAAGCAATGATTAAGCAAACTCAA 480
Db 23825 ACAGTGAAGTATCTGCAAGTTTGAATGTAACAAAGCAATGATTAAGCAAACTCAA 23766
QY 481 TGGATTACTTAACAAAGAGTTAAACAGATGCTATAAGTTAAGACATGACACGACGA 540
Db 23765 TGGATTACTTAACAAAGAGTTAAACAGATGCTATAAGTTAAGACATGACACGACGA 23706
QY 541 GTGTTAACTGGAAGCGTAGAGTTAATTAAGTATCCACAATATACACAGTATCAA 600
Db 23705 GTGTTAACTGGAAGCGTAGAGTTAATTAAGTATCCACAATATACACAGTATCAA 23646
QY 601 GATGGAAGAAACCCATATAAGTAAATTTGCTAAGCTGGGAATATTCATTTATACAG 660
Db 23645 GATGGAAGAAACCCATATAAGTAAATTTGCTAAGCTGGGAATATTCATTTATACAG 23586
QY 661 CAAATTTATCAAGATGAATCAACACCAATTTGATATATCTGAAATCCAAATTTAACA 720
Db 23585 CAAATTTATCAAGATGAATCAACACCAATTTGATATATCTGAAATCCAAATTTAACA 23526
QY 721 AGCAATCAACACATATGAAATTTACTTGAAGCTGTGATTTAACAACAATTCCT 780
Db 23525 AGCAATCAACACATATGAAATTTACTTGAAGCTGTGATTTAACAACAATTCCT 23466
QY 781 CACAATTTATTAAGAGAGAGTTGAACTTGTAGTGAAGGAGTCCAACTACT 840
Db 23465 CACAATTTATTAAGAGAGAGTTGAACTTGTAGTGAAGGAGTCCAACTACT 23406
QY 841 TCAGAAATCTCCATCT--ATTATATAGATGAATTTATCTATTATGCCCGAAATTA 897
Db 23405 TCAGAAATCTCCATCTAAATATATAGATGAATTTATCTATTATGCCCGAAATTA 23346
QY 898 ATTGGCGGATCTGCAAAATTTATCAATTAACAACAATGATGATGAGTACCAGAT 957
Db 23345 ATTGGCGGATCTGCAAAATTTATCAATTAACAACAATGATGATGAGTACCAGAT 23286
QY 958 GCGAACCATTGAAATGTCATTCGAGTTATTAATCAAAATGTTAAATTAACCTTA 1017
Db 23285 GCGAACCATTGAAATGTCATTCGAGTTATTAATCAAAATGTTAAATTAACCTTA 23226
QY 1018 CGAAAGAGTGA 1029
Db 23225 CGAAAGAGTGA 23214
```

Search completed: November 19, 2000, 05:56:05
Job time: 8392 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 05:56:05 : Search time 2066.03 Seconds
(without alignments)
1263.820 Million cell updates/sec

Title: US-08-978-456-3
Perfect score: 1269
Sequence: 1 AANCACCAATCCNATGCGA.....GTAAACAAGATGCTATAG 1269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10041192 seqs, 1028798125 residues

Total number of hits satisfying chosen parameters: 20082384

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Pending_Patents_NA:*
1: /cgn2_6/ptodata/2/pna/US0822C.COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06.COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07.COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08.COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US081A.COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US081B.COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US081C.COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US081D.COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US082A.COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US082B.COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US082C.COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US083A.COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US083B.COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US084A.COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US084B.COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US084C.COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US085.COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US087A.COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US087B.COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US087C.COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US088A.COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US088B.COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US088C.COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US088D.COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US089A.COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US089B.COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US089C.COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US089D.COMB.seq:*
29: /cgn2_6/ptodata/2/pna/US089E.COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US089F.COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US089G.COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US089H.COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US089I.COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US090A.COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US090B.COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US090C.COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US090D.COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US090E.COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US091A.COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US091B.COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US091C.COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US092A.COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US092B.COMB.seq:*

44: /cgn2_6/ptodata/2/pna/US092B.COMB.seq:*
45: /cgn2_6/ptodata/2/pna/US092C.COMB.seq:*
46: /cgn2_6/ptodata/2/pna/US092D.COMB.seq:*
47: /cgn2_6/ptodata/2/pna/US092E.COMB.seq:*
48: /cgn2_6/ptodata/2/pna/US093A.COMB.seq:*
49: /cgn2_6/ptodata/2/pna/US093B.COMB.seq:*
50: /cgn2_6/ptodata/2/pna/US093C.COMB.seq:*
51: /cgn2_6/ptodata/2/pna/US093D.COMB.seq:*
52: /cgn2_6/ptodata/2/pna/US093E.COMB.seq:*
53: /cgn2_6/ptodata/2/pna/US093F.COMB.seq:*
54: /cgn2_6/ptodata/2/pna/US094A.COMB.seq:*
55: /cgn2_6/ptodata/2/pna/US094B.COMB.seq:*
56: /cgn2_6/ptodata/2/pna/US094C.COMB.seq:*
57: /cgn2_6/ptodata/2/pna/US095.COMB.seq:*
58: /cgn2_6/ptodata/2/pna/US096.COMB.seq:*
59: /cgn2_6/ptodata/2/pna/US097A.COMB.seq:*
60: /cgn2_6/ptodata/2/pna/US097B.COMB.seq:*
61: /cgn2_6/ptodata/2/pna/US098A.COMB.seq:*
62: /cgn2_6/ptodata/2/pna/US098B.COMB.seq:*
63: /cgn2_6/ptodata/2/pna/US098C.COMB.seq:*
64: /cgn2_6/ptodata/2/pna/US098D.COMB.seq:*
65: /cgn2_6/ptodata/2/pna/US098E.COMB.seq:*
66: /cgn2_6/ptodata/2/pna/US098F.COMB.seq:*
67: /cgn2_6/ptodata/2/pna/US098G.COMB.seq:*
68: /cgn2_6/ptodata/2/pna/US098H.COMB.seq:*
69: /cgn2_6/ptodata/2/pna/US098I.COMB.seq:*
70: /cgn2_6/ptodata/2/pna/US099A.COMB.seq:*
71: /cgn2_6/ptodata/2/pna/US099B.COMB.seq:*
72: /cgn2_6/ptodata/2/pna/US099C.COMB.seq:*
73: /cgn2_6/ptodata/2/pna/US099D.COMB.seq:*
74: /cgn2_6/ptodata/2/pna/US099E.COMB.seq:*
75: /cgn2_6/ptodata/2/pna/US099F.COMB.seq:*
76: /cgn2_6/ptodata/2/pna/US099G.COMB.seq:*
77: /cgn2_6/ptodata/2/pna/US099H.COMB.seq:*
78: /cgn2_6/ptodata/2/pna/US099I.COMB.seq:*
79: /cgn2_6/ptodata/2/pna/US099J.COMB.seq:*
80: /cgn2_6/ptodata/2/pna/US099K.COMB.seq:*
81: /cgn2_6/ptodata/2/pna/US099L.COMB.seq:*
82: /cgn2_6/ptodata/2/pna/US099M.COMB.seq:*
83: /cgn2_6/ptodata/2/pna/US099N.COMB.seq:*
84: /cgn2_6/ptodata/2/pna/US099O.COMB.seq:*
85: /cgn2_6/ptodata/2/pna/US099P.COMB.seq:*
86: /cgn2_6/ptodata/2/pna/US099Q.COMB.seq:*
87: /cgn2_6/ptodata/2/pna/US099R.COMB.seq:*
88: /cgn2_6/ptodata/2/pna/US099S.COMB.seq:*
89: /cgn2_6/ptodata/2/pna/US099T.COMB.seq:*
90: /cgn2_6/ptodata/2/pna/US099U.COMB.seq:*
91: /cgn2_6/ptodata/2/pna/US099V.COMB.seq:*
92: /cgn2_6/ptodata/2/pna/US099W.COMB.seq:*
93: /cgn2_6/ptodata/2/pna/US099X.COMB.seq:*
94: /cgn2_6/ptodata/2/pna/US099Y.COMB.seq:*
95: /cgn2_6/ptodata/2/pna/US099Z.COMB.seq:*
96: /cgn2_6/ptodata/2/pna/US099A.COMB.seq:*
97: /cgn2_6/ptodata/2/pna/US099B.COMB.seq:*
98: /cgn2_6/ptodata/2/pna/US099C.COMB.seq:*
99: /cgn2_6/ptodata/2/pna/US099D.COMB.seq:*
100: /cgn2_6/ptodata/2/pna/US099E.COMB.seq:*
101: /cgn2_6/ptodata/2/pna/US099F.COMB.seq:*
102: /cgn2_6/ptodata/2/pna/US099G.COMB.seq:*
103: /cgn2_6/ptodata/2/pna/US099H.COMB.seq:*
104: /cgn2_6/ptodata/2/pna/US099I.COMB.seq:*
105: /cgn2_6/ptodata/2/pna/US099J.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1250	98.5	1269	1	PCR-US97-02318-946 Sequence 946, App


```
|||||
Db 601 CGCAAGATATGATGCGATTTGTAATAATGTACAAATAGCGTTATTATTAACATTA 660
QY 661 ATTTTCTCCTTGCATCTTAATTCATGATGAGAGATTTTGTATTAGAGGTATCA 720
Db 661 ATTTTCTCCTTGCATCTTAATTCATGATGAGATTTTGTATTAGAGGTATCA 720
QY 721 TTGTAGTCAATTAATGATTAATGCGATTCACCTCAAAATAGTGTCAAGGTACAGG 780
Db 721 TTGTAGTCAATTAATGATTAATGCGATTCACCTCAAAATAGTGTCAAGGTACAGG 780
QY 781 TGTATTCACCGCTGGCGCTTGTAGTTAATGAAGTAGAGATTGTGGTATGGTGC 840
Db 781 TGTATTCACCGCTGGCGCTTGTAGTTAATGAAGTAGAGATTGTGGTATGGTGC 840
QY 841 ACACCTTGAGAAAAAGTGACACAGCATCGGAGGTTCAAGCACTTGATATGGCACACANAA 900
Db 841 ACACCTTGAGAAAAAGTGACACAGCATCGGAGGTTCAAGCACTTGATATGGCACACANAA 900
QY 901 TGGTGAAGGTGGCAGATTTATTTAGTTAGTAGAGCATGTAGTCATTTGGTTCACACACC 960
Db 901 TGGTGAAGGTGGCAGATTTATTTAGTTAGTAGAGCATGTAGTCATTTGGTTCACACACC 960
QY 961 ACCCTGTGTTAACAAATTAATGATTTGAAGTAGACANAAAGTAGTATTACNCAACANAAAG 1020
Db 961 ACCCTGTGTTAACAAATTAATGATTTGAAGTAGACANAAAGTAGTATTACNCAACANAAAG 1020
QY 1021 ACAATTCCTGTAGACACACATGGGTATGACAGCTTACGGGGCTCCACGGTATTGAGGG 1080
Db 1021 ACAATTCCTGTAGACACACATGGGTATGACAGCTTACGGGGCTCCACGGTATTGAGGG 1080
QY 1081 TTGAATTTGGTGGATGATGAAGGGCATGCACAAATTAACCAAGACTTTTAAAGCA 1140
Db 1081 TTGAATTTGGTGGATGATGAAGGGCATGCACAAATTAACCAAGACTTTTAAAGCA 1140
QY 1141 AAAAGCAAGCAACTTGCACAAATTTACAGTAAGTNTCTTGAAGTTTATGATGGTGA 1200
Db 1141 AAAAGCAAGCAACTTGCACAAATTTACAGTAAGTNTCTTGAAGTTTATGATGGTGA 1200
QY 1201 AACAAAGCAATTAATGATGAAGACAAAGTCAATGATTAACAAAGAGTTAAACAGA 1260
Db 1201 AACAAAGCAATTAATGATGAAGACAAAGTCAATGATTAACAAAGAGTTAAACAGA 1260
QY 1261 TGTCTATAG 1269
Db 1261 TGTCTATAG 1269

RESULT 2
PCT-US98-24857-3
; Sequence 3 Application PC/TUS9824857A
; GENERAL INFORMATION:
; APPLICANT: Smithline Beecham Corporation
; TITLE OF INVENTION: Novel ribd
; FILE REFERENCE: P50444-9
; CURRENT APPLICATION NUMBER: PCT/US98/24857A
; EARLIER APPLICATION NUMBER: 08/978,456
; EARLIER FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1269
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1269)
; OTHER INFORMATION: n = A,T,C or G
PCT-US98-24857-3
```

Query Match

98.5%; Score 1250; DB 1; Length 1269;

```
Best Local Similarity 100.0%; Pred. No. 4,76-269;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AANCACCAATCCNATTGGGAGGNAATCCAAATCAATNCCCGGANNCCCATCCAGTTAA 60
Db 1 AANCACCAATCCNATTGGGAGGNAATCCAAATCAATNCCCGGANNCCCATCCAGTTAA 60
QY 61 TTAAGTCCAGGTTTGGAGCAATTAACCAATATGATTCGATGAGGTCAATGNCANAG 120
Db 61 TTAAGTCCAGGTTTGGAGCAATTAACCAATATGATTCGATGAGGTCAATGNCANAG 120
QY 121 GTGTTATTAACATACGAATGNTGTGNAATATAGTAGANCAAGTTGGCATACAGTAT 180
Db 121 GTGTTATTAACATACGAATGNTGTGNAATATAGTAGANCAAGTTGGCATACAGTAT 180
QY 181 CTGCAATTAATGATGCTTTAGAAATGCAATGACACACAGCGCAGTTCAATCAATAT 240
Db 181 CTGCAATTAATGATGCTTTAGAAATGCAATGACACACAGCGCAGTTCAATCAATAT 240
QY 241 TTGTACTAGAGATTAATAAGAAAGCGCTATAGAGCAATGGAAGTTGATTTTA 300
Db 241 TTGTACTAGAGATTAATAAGAAAGCGCTATAGAGCAATGGAAGTTGATTTTA 300
QY 301 ATGCTGTAGTAGAATCATATCATGAGATGCCATATAGTACTGAGATTATTAATTAAT 360
Db 301 ATGCTGTAGTAGAATCATATCATGAGATGCCATATAGTACTGAGATTATTAATTAAT 360
QY 361 AAAACCGTATTAATGTTTTTTTACAAAACATATAGTATCAATTTTAAATGATTTGACA 420
Db 361 AAAACCGTATTAATGTTTTTTTACAAAACATATAGTATCAATTTTAAATGATTTGACA 420
QY 421 TACTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 TACTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 481 ATTCCCAACCGGAGTAAATTAAGCGCTGACCTGCTAATATGTTTCAATTTAGTGGCTG 540
Db 481 ATTCCCAACCGGAGTAAATTAAGCGCTGACCTGCTAATATGTTTCAATTTAGTGGCTG 540
QY 541 ATCTAGTGAATTTAGACCCGACAGTAAAGTCTGATGAGAGAAAGATTAATTAAT 600
Db 541 ATCTAGTGAATTTAGACCCGACAGTAAAGTCTGATGAGAGAAAGATTAATTAAT 600
QY 601 CGCAAGATTAATGATGAGCTATTTGTAATAATGTGACAAATAGGCTTTTAAACGATA 660
Db 601 CGCAAGATTAATGATGAGCTATTTGTAATAATGTGACAAATAGGCTTTTAAACGATA 660
QY 661 ATTTTCTCCTTGCATCTTAATTCATGATGAGAGATTTTGTATTAGAGGTATCA 720
Db 661 ATTTTCTCCTTGCATCTTAATTCATGATGAGAGATTTTGTATTAGAGGTATCA 720
QY 721 TTGTAGTCAATTAATGATTAATGCGATTCACCTCAAAATAGTGTCAAGGTACAGG 780
Db 721 TTGTAGTCAATTAATGATTAATGCGATTCACCTCAAAATAGTGTCAAGGTACAGG 780
QY 781 TGTATTCACCGCTGGCGCTTGTAGTTAATGAAGTAGAGATTGTGGTATGGTGC 840
Db 781 TGTATTCACCGCTGGCGCTTGTAGTTAATGAAGTAGAGATTGTGGTATGGTGC 840
QY 841 ACACCTTGAGAAAAAGTGACACAGCATCGGAGGTTCAAGCACTTGATATGGCACACANAA 900
Db 841 ACACCTTGAGAAAAAGTGACACAGCATCGGAGGTTCAAGCACTTGATATGGCACACANAA 900
QY 901 TGGTGAAGGTGGCAGATTTATTTAGTTAGTAGAGCATGTAGTCATTTGGTTCACACACC 960
Db 901 TGGTGAAGGTGGCAGATTTATTTAGTTAGTAGAGCATGTAGTCATTTGGTTCACACACC 960
QY 961 ACCCTGTGTTAACAAATTAATGATTTGAAGTAGACANAAAGTAGTATTACNCAACANAAAG 1020
Db 961 ACCCTGTGTTAACAAATTAATGATTTGAAGTAGACANAAAGTAGTATTACNCAACANAAAG 1020
QY 1021 ACAATTCCTGTAGACACACATGGGTATGAGAGCTTACGGGGCTCCACGGTATTGAGGG 1080
Db 1021 ACAATTCCTGTAGACACACATGGGTATGAGAGCTTACGGGGCTCCACGGTATTGAGGG 1080
```

Db 1021 acaatccgttagacacacatcgggtgatgagacgttacccgggctccacgctatttgaggg 1080
QY 1081 TTGAATTCGCTTGATGATGAACGGGCAATCAATTTATCCAAAGACTTTTTRAAACA 1140
Db 1081 ttgaatttcgcttgatgatgaaacgggcatcacaattatccaaagacttttttaaagca 1140
QY 1141 AAAAGCAAGCAACTTGGCACAATAATACAGTGAAGTCTTGAAGATTAGATGGTA 1200
Db 1141 aaagcaagaacttcgcacacaaattcagtgaaagtcttgaagattcagatggta 1200
QY 1201 AACAAAGCAATTTGATTAATGAACAAGTCAATGATTTACTTAACAAGGTTAAACAAGA 1260
Db 1201 aacaagaagattgataatgacaaagcgaatgatacttaacaaagaggttaacaaga 1260
QY 1261 TGTCTATAG 1269
Db 1261 tgcctatag 1269

RESULT 3
US-08-903-470-946
; Sequence 946, Application US/08903470
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Raymond
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: Novel Compounds
; NUMBER OF SEQUENCES: 1166
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/903,470
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/011,888
; FILING DATE: 20-FEB-1996
; APPLICATION NUMBER: PCT/US97/02318
; FILING DATE: 19-FEB-1997
; APPLICATION NUMBER: PCT/US97/02547
; FILING DATE: 19-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimm, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: GM50006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 946:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1269 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-903-470-946

Query Match 98.5%; Score 1250; DB 26; Length 1269;
Best Local Similarity 100.0%; Pred.No. 4,7e-265;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANACCAATCCNATTTGGAGNATCCAAATCAATNCCCGGANCCCAATCCAAAGTTAA 60
Db 1 AANACCAATCCNATTTGGAGNATCCAAATCAATNCCCGGANCCCAATCCAAAGTTAA 60
QY 61 TTAAGTCAAGGTTTGGAAACATTACCAATATGATTTCCGATAGGTCAATGNCANCG 120
Db 61 TTAAGTCAAGGTTTGGAAACATTACCAATATGATTTCCGATAGGTCAATGNCANCG 120
QY 121 GTGTTAATAACTACGAATGNTGTGNAATGATAGAGNCAAGTTGGCCATACAGTNT 180
Db 121 GTGTTAATAACTACGAATGNTGTGNAATGATAGAGNCAAGTTGGCCATACAGTNT 180
QY 181 CTCNATTATATGATGCTTTAGAAATCGAATGAGCACAACAGCCAGTTACCAATCAATAT 240
Db 181 CTCNATTATATGATGCTTTAGAAATCGAATGAGCACAACAGCCAGTTACCAATCAATAT 240
QY 241 TTGTAAGTACAGATTAATAAGAGACGCTCTATAGAGCAATTCAGGTTGATTTTA 300
Db 241 TTGTAAGTACAGATTAATAAGAGACGCTCTATAGAGCAATTCAGGTTGATTTTA 300
QY 301 ATGCTGTTAGTAGATCATCATCAATGAGATGCTATAGTACTCAGATTAATTAAT 360
Db 301 ATGCTGTTAGTAGATCATCATCAATGAGATGCTATAGTACTCAGATTAATTAAT 360
QY 361 AAAACCGTCATTAATGTTTTTTTGAAGAACATATAGTATCATTTTAAATGATGACA 420
Db 361 AAAACCGTCATTAATGTTTTTTTGAAGAACATATAGTATCATTTTAAATGATGACA 420
QY 421 TACTAGTACTCAATATATCTATACCAATTTCAATATTAATCTTTGGGGCAGGGTGA 480
Db 421 TACTAGTACTCAATATATCTATACCAATTTCAATATTAATCTTTGGGGCAGGGTGA 480
QY 481 ATTCCCAACGGCGAGTAATAAAGCCCTGCGACCTGCTAATATGTTTCAATTTAGTGCTG 540
Db 481 ATTCCCAACGGCGAGTAATAAAGCCCTGCGACCTGCTAATATGTTTCAATTTAGTGCTG 540
QY 541 ATCTAGTACATTTCTAGAGCCGACAGTTAAAGTCTGATGGAGAAAGATGTTAATAT 600
Db 541 ATCTAGTACATTTCTAGAGCCGACAGTTAAAGTCTGATGGAGAAAGATGTTAATAT 600
QY 601 CGACAAAGATATATGATCGATTTGTAAATGTGTACCAATATGCGTTATTTAACGTAA 660
Db 601 CGACAAAGATATATGATCGATTTGTAAATGTGTACCAATATGCGTTATTTAACGTAA 660
QY 661 ATTTTCTCCTTGCATCTTAATTCATGATGTGAGATTTTGTGTTATAGAGTGATCA 720
Db 661 ATTTTCTCCTTGCATCTTAATTCATGATGTGAGATTTTGTGTTATAGAGTGATCA 720
QY 721 TTTGAGTCAATTTATGATTAATGATGATTCACCTTCCAAATATGTTACAGTCAAGG 780
Db 721 TTTGAGTCAATTTATGATTAATGATGATTCACCTTCCAAATATGTTACAGTCAAGG 780
QY 781 TGTTAATCCACCGCTGGCGCTGTGTAGTTAATGAAGTGAAGATTTGTGTTGTGTC 840
Db 781 TGTTAATCCACCGCTGGCGCTGTGTAGTTAATGAAGTGAAGATTTGTGTTGTGTC 840
QY 841 ACACCTTGAGAAAGGTGACAAAGCATGCGGAGTTCAAGCACTTGATATGCGACAACANAA 900
Db 841 ACACCTTGAGAAAGGTGACAAAGCATGCGGAGTTCAAGCACTTGATATGCGACAACANAA 900
QY 901 TGCTGAAGGTGCGACGATTAATATTTAGCTAGAGCCATGAGTCAATTTTGGTCAACACC 960
Db 901 TGCTGAAGGTGCGACGATTAATATTTAGCTAGAGCCATGAGTCAATTTTGGTCAACACC 960
QY 961 ACCCTGTGTTAACAAATATTTGATTTGATAGATAGCANAAAGTATTTACNCAACANAAG 1020
Db 961 ACCCTGTGTTAACAAATATTTGATTTGATAGATAGCANAAAGTATTTACNCAACANAAG 1020


```
|||||
Db 1141 AAAGCAAGCAACTTGCACAAAATACAGTGAAGTCTTGAAGTTAGATGGGTA 1200
Qy 1201 AACAAAGGATTTGATATGACAAAGTCATGATTTACTAACAAAGGTTAAACAGA 1260
Db 1201 AACAAAGGAAATTTGATATGACAAAGTCATGATTTACTAACAAAGGTTAAACAGA 1260
Qy 1261 TGTCTATAG 1269
Db 1261 TGTCTATAG 1269

RESULT 5
US-09-611-529-7446/c
; Sequence 7446, Application US/09611529
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Roberta S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
; FILE REFERENCE: 1034/1C963US1
; CURRENT APPLICATION NUMBER: US/09/611,529
; PRIOR APPLICATION NUMBER: US 09/417,811
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/353,718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,334
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,221
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,137
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,082
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,081
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,079
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/035,913
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/035,744
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 08/827,356
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: US 08/831,156
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: US 60/014,477
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: US 60/016,743
; PRIOR FILING DATE: 1996-05-02
; PRIOR APPLICATION NUMBER: US 60/020,016
; PRIOR FILING DATE: 1996-06-14
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 7446
; LENGTH: 74105
```

```
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; US-09-611-529-7446

Query Match 76.1%; Score 965.6; DB 57; Length 74105;
Best Local Similarity 93.5%; Pred. No. 5e-205;
Matches 1123; Conservative 0; Mismatches 57; Indels 21; Gaps 11;

Qy 68 CAAGTTTGGACATTCACCAATATGATCCGATGGATGCAATGCAATGCGGTAA 127
Db 24908 CCAAGTTTGGACATTCACCAATATGATCCGAT-AGGTCAAATGCAAAAGGTGTTAA 24850
Qy 128 TAACTACGAATGNTGTGNAATAGATAGTACAGCAAGTTGCCATACAGTCTCAAT 187
Db 24849 TAACTACGAATGNTGTGNAATAGATAGTACAGCAAGTTGCCATACAGTCTCAAT 24790
Qy 188 ATATGATGCTTTAGATGCAATGCAACCAACACCGAGTTACATCAATATTTGTAAC 247
Db 24789 ATATGATGCTTTAGATGCAATGCAACCAACACCGAGTTACATCAATATTTGTAAC 24730
Qy 248 TAGAAGATTAATAAGAGAGCGCTCTATAGAGAGATGGAAGTTGATTTATGCTG 307
Db 24729 TAGAAGATTAATAAGAGAGCGCTCTATAGAGAGATGGAAGTTGATTTATGCTG 24670
Qy 308 TTAGTAAATCATATCATATGAGATGCGCTATAGTACTCAGATTATTAATTAATTAACCG 367
Db 24669 TTAGTAAATCATATCATATGAGATGCGCTATAGTACTCAGATTATTAATTAATTAACCG 24610
Qy 368 TCATTATTTGTTTTTTAGAAACATATAGTATCATTTTAAATGATGACATCTACG 427
Db 24609 TCATTATTTGTTTTTTAGAAACATATAGTATCATTTTAAATGATGACATCTACG 24550
Qy 428 TACTCAATATATCTATTAACATTTTCATATATTTCTTGGGGCAGGTAATCCCA 487
Db 24549 TACTCAATATATCTATTAACATTTTCATATATTTCTTGGGGCAGGTAATCCCA 24490
Qy 488 ACCGCGATTAATAAGACCTGCGACCTGCTAATATGTTTCATATTTAGTGGCTGATCTAGT 547
Db 24489 ACCGCGATTAATAAGACCTGCGACCTGCTAATATGTTTCATATTTAGTGGCTGATCTAGT 24430
Qy 548 GAGATTCTAGACCCGACAGTTAAAGTCTGATGGGAGCAAAAGATTTATGACAAA 607
Db 24429 GAGATTCTAGACCCGACAGTTAAAGTCTGATGGGAGCAAAAGATTTATGACAAA 24370
Qy 608 GATATATGAGCGTATTTGTAAATGTTGTAATAGGCTTATTAACGATTAATTTTC 667
Db 24369 GATATATGAGCGTATTTGTAAATGTTGTAATAGGCTTATTAACGATTAATTTTC 24310
Qy 668 TCCTTGATCTTATATCATGATGAGGATTTTTTTGTATAGAGGTGATCATTTGAGT 727
Db 24309 TCC-TTGATCTTATATCATGATGAGGAT-TTTTTGTATAGAGGTGATCATTTGAGT 24252
Qy 728 CAATTTATGATTTATGCGATTCAACTTCCAAATATGTTGTAAGGTGATTAAT 787
Db 24251 CAATTTATGATTTATGCGATTCAACTTCCAAATATGTTGTAAGGTGATTAAT 24192
Qy 788 CCACCCGTTGGCGCTGTTGTATGATGAAGGTGAGATTGTTGTTATGTCACACTTG 847
Db 24191 CCACCCGTTGGCGCTGTTGTATGATGAAGGTGAGATTGTTGTTATGTCACACTTG 24132
Qy 848 AGAAAGGTGACAGCATGCGAGGTTCAAGCACTTATGACCAACAAATGCTGAA 907
Db 24131 AGAAAGGTGACAGCATGCGAGGTTCAAGCACTTATGACCAACAAATGCTGAA 24072
Qy 908 GGTGCGAGATTTATATTTAGCTTAGACCATGTAGTATTTGTTTAAACACCCCTGT 967
Db 24071 GGTGCGAGATTTATATTTAGCTTAGACCATGTAGTATTTGTTTAAACACCCCTGT 24012
Qy 968 GTTAACAAATATTTATGATTTGAATAGTACCAATAGTATTAACCAACAAAGCAATT 1027
Db 24011 GTTAACAAATATTTATGATTTGAATAGTACCAATAGTATTAACCAACAAAGCAATT- 23952
```



```

QY      1028 CGTGTGACACACATGGTGATAGACGTTACGGGGCTCCAGCGTATTTAGGGTTGAATT 1087
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       23953 CGTTTGACACACAT -GGTGATGAGAGCTTACGGGCTCACGGTAATGAGTTGCATT 23895
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1088 GCGTTGATGATGACAGCGGCATCCACATATATACCAAGAAGCTTTTTTAAAGCAAAGC 1147
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       23894 -----GAGGATGAMCGGGCATCCACAATTATACCA---AGACTTTTTTAAAGCAAAGC 23844
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1148 AAGCAACTTGGCCACAAAATTTAACAGTGAAGAGTNTCTTGAAGATTAGATGGGTTAACAAG 1207
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       23843 AAGCAACT--GCCACAAATTTACAGTGAAGAGTATC-TGCAAGTTTACATGTTAAACAAGC 23788
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1208 CCAATTGATTAATGACAAAGTCATGATGATTACTAACAAAGAGGTTAAACAAGTGTCTAT 1267
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       23786 ---AATGATAAATGAGCAAAAGTCATGATTACTTAACAAAGAGGTTAAACAAGATGTCTAT 23730
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1268 A 1268
          |
Db       23729 A 23729
          |

RESULT      6
US-08-831-156A-107/C
; Sequence 107 Application US/08831156A
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Roberta S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE OF A
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS
; TITLE OF INVENTION: GENOME AND RELATED
; NUMBER OF SEQUENCES: COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 112
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corporation
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033-0530

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/831.156A
FILING DATE: 01-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/014,477
FILING DATE: 01-APR-1996
APPLICATION NUMBER: 60/016,743
FILING DATE: 02-MAY-1996
APPLICATION NUMBER: 60/020,016
FILING DATE: 14-JUN-1996
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 85137 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Staphylococcus aureus
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...85161

```

Query Match	76.1%	Score 965.6	DB 23	Length 85137
Best Local Similarity	93.5%	Pred. No. 5.2e-203		
Matches 1123	Conservative	0	Mismatches 57	Indels 21
				Gaps 11
QY	68	CAGGCTTTGGAAACATTTACCAATATGATCCGATGAGGTCCAATGNCANCGGTGTAA	127	
Db	50771	CCAAGCTTTGGAAACATTTAACAATATGATGGAT-AGGTCAAATGACAAACGGTGTAA	50713	
QY	128	TAACTACGAATGNTGTGMAATGATAGTAGANCAAGTTGGCATACAGTCTCMAAT	187	
Db	50712	TAACTACGAATGNTGTGMAATGATAGTAGANCAAGTTGGCATACAGTCTCMAAT	50653	
QY	188	ATATGATGCTTTAGATGCAATGAGCACACACAGCGCACTTCAATCATATTTGTAC	247	
Db	50652	ATATGATGCTTTAGATGCAATGAGCACACACAGCGCACTTCAATCATATTTGTAC	50593	
QY	248	TGAAGATATATAAGAAACGCTCTATAGACGAATTCAGCTTGATTTAATGTG	307	
Db	50592	TGAAGATATATAAGAAACGCTCTATAGACGAATTCAGCTTGATTTAATGTG	50533	
QY	308	TTAGTAGAATCATATCAATGAGATGCTATAGTACTGATTAATATTTAAATAAACG	367	
Db	50532	TTAGTAGAATCATATCAATGAGATGCTATAGTACTGATTAATATTTAAATAAACG	50473	
QY	368	TCATTAATTTGTTTTTTAGAAAACATATAGTATCATTTTAAATGTATAGTACATCTACG	427	
Db	50472	TCATTAATTTGTTTTTTAGAAAACATATAGTATCATTTTAAATGTATAGTACATCTACG	50413	
QY	428	TACTCAATATATCTATTAACAATTTATATATATATTTCTTTGGGGGACGGGTGAATCCCA	487	
Db	50412	TACTCAATATATCTATTAACAATTTATATATATATTTCTTTGGGGGACGGGTGAATCCCA	50353	
QY	488	ACCGGCAGTAAATAAAGCCTGCGACCTGCTAATATGTTTCATATTAAGTGGCTGATCTAGT	547	
Db	50352	ACCGGCAGTAAATAAAGCCTGCGACCTGCTAATATGTTTCATATTAAGTGGCTGATCTAGT	50293	
QY	548	GAGATTTAGAGCCGACACAGTTAAAGTCTGAGTGGGAGAAAGATGTATATTCGACAA	607	
Db	50292	GAGATTTAGAGCCGACACAGTTAAAGTCTGAGTGGGAGAAAGATGTATATTCGACAA	50233	
QY	608	GATTAATGATGCTATTTGTAAAAATGTGACAAATAGGCTTATTAACGATTAATTTTC	667	
Db	50232	GATTAATGATGCTATTTGTAAAAATGTGACAAATAGGCTTATTAACGATTAATTTTC	50173	
QY	668	TGCTTTGCATCTTAATTCATGATGTGAGGATTTTTGTTTATAGAGGTGATCATTTGACT	727	
Db	50172	TGCTTTGCATCTTAATTCATGATGTGAGGATTTTTGTTTATAGAGGTGATCATTTGACT	50115	
QY	728	CAATTTATGATTTATGCGATTCACCTCCAAATATGTACAAAGTTCANACAGGTGTAA	787	
Db	50114	CAATTTATGATTTATGCGATTCACCTCCAAATATGTGTACAAGTTCANACAGGTGTAA	50055	
QY	788	CCACCCGTTGGCGCTGTGTAGTTAATGAAGTAGATGTGTGTATGCTGCACACTTG	847	
Db	50054	CCACCCGTTGGCGCTGTGTATTAATGAAGTAGAGATGTGTGTATGCTGCACACTTG	49995	
QY	848	AGAAAGGTGACAAACATGCGAGGTTCAAGCACTTGATATGGCAACANATGCTGA	907	
Db	49994	AGAAAGGTGACAAACATGCGAGGTTCAAGCACTTGATATGGCAACANATGCTGA	49935	
QY	908	GGTGGCAGATTTATATTACGTTAGAGCCATGTAGTCAATTTTGGTTCAACACCCTGT	967	
Db	49934	GGTGGCAGATTTATATTACGTTAGAGCCATGTAGTCAATTTTGGTTCAACACCCTGT	49875	
QY	968	GTTAACAAATATATGATTTAGATGTAGATACANAAAGTATTTACNCAACANAAACAATT	1027	
Db	49874	GTTAACAAATATATGATTTAGATGTAGATACANAAAGTATTTACNCAACANAAACAATT	49815	
QY	1028	CGTTAGACACACATGGGTGATGAGACGTTACGGGGCTCCACAGGTATTTGAGGGTGAAT	1087	
Db	49816	CGTTAGACACACAT-GGGTGATGAGACGTTACGGGGCTCCACAGGTATTTGAGGGTGAAT	49758	
QY	1088	GGGTGGATGATGAACGGGCAATCACAATTTATCCAAAGACTTTTTTTAAAGCAAAAAGA	1147	


```
RESULT
US-08-956-171-142/C
; Sequence 142, Application US/08956171
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 142:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7588 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-08-956-171-142

Query Match          75.5%  Score 957.6:  DB 30:  Length 7588;
Best Local Similarity 93.1%  Pred. No. 1.6e-203;
Matches 1118; Conservative 0; Mismatches 62; Indels 21; Gaps 11;
```

```

Db 6153 TACTCAAAATATCTATACAAATTTTCATATATATCTTTTCGGGGCAGGGTGAATTCCTCA 6094
QY 488 ACCGCGAGTAATAAAGCGCTGCAGCTGCTAATATGTTTCATATATGTCGATCTAGT 547
Db 6093 ACCGCGAGTAATAAAGCGCTGCAGCTGCTAATATGTTTCATATATGTCGATCTAGT 6034
QY 548 GAGATCTGAGCGCAGCTGTAAGAGCTGAGAGGAGAGAAATGTAATTTATCGCAAA 607
Db 6033 GAGATCTGAGCGCAGCTGTAAGAGCTGAGAGGAGAGAAATGTAATTTATCGCAAA 5974
QY 608 GATAATGACGCTATTGTTAAAGATGTACAAATAGCGCTTATTATACATAATTTTC 667
Db 5973 GATAATGACGCTATTGTTAAAGATGTACAAATAGCGCTTATTATACATAATTTTC 5914
QY 668 TCCTTTGCACTTAAATTCATGATGAGATTTTGTATTATAGCTGATCATTTGAGT 727
Db 5913 TCC-TTGCATCTTAAATTCATGATGATGAGA-TTTTGTATTATAGCTGATCATTTGAGT 5856
QY 728 CAATTATGATTTATGCGATTCACCTTCCAAATATGTGTACAAAGTCCANACAGCTGTAAT 787
Db 5855 CAATTATGATTTATGCGATTCACCTTCCAAATATGTGTACAAAGTCCANACAGCTGTAAT 5796
QY 788 CCACCCGTTGGCGCTGTTAGTTAATGAAGTAGAGATTGTGTATTGTCACACTTG 847
Db 5795 CCACCCGTTGGCGCTGTTAGTTAATGAAGTAGAGATTGTGTATTGTCACACTTG 5736
QY 848 AGAAAGGTGACAGCATCGGAGGTTCAAGCACTTGATATGCGCACAAANATGCTGAA 907
Db 5735 AGAAAGGTGACAGCATCGGAGGTTCAAGCACTTGATATGCGCACAAANATGCTGAA 5676
QY 908 GGTGCGACGATTTATATTAAGTTAGAGCATGATGATTTGTTGTTCAACACACCTGCT 967
Db 5675 GGTGCGACGATTTATATTAAGTTAGAGCATGATGATTTGTTGTTCAACACACCTGCT 5616
QY 968 GTTAACAAATATTATGTTGTAAGATGACAAAGTAGATTTACNCACAAANAGACAATTC 1027
Db 5615 GTTAACAAATATTATGTTGTAAGATGACAAAGTAGATTTACNCACAAANAGACAATTC 5556
QY 1028 CGTTAGACACATGCGGTGATGAGACGTTACGGGCGCTCCACGGTATTGAGGTTGAATT 1087
Db 5557 CGTTAGACACATGCGGTGATGAGACGTTACGGGCGCTCCACGGTATTGAGGTTGAATT 5499
QY 1088 GCGTTGATGATGAACGGGCTATCACAATTTATACAAAGACTTTTATTAAGCAAAAAGCA 1147
Db 5498 -----GATGATGAACGGGCTATCACAATTTATACCA---AGACTTTTATTAAGCAAAAAGCA 5448
QY 1148 AAGCACTTGGCCACAAATTTACAGTGAAGTCTTTGAAGTTTACATGGTTAAACCAAG 1207
Db 5447 AAGCACT--GCCACAAATTTACAGTGAAGTATC--TGCAAGTTTACATGGTTAAACCAAG 5391
QY 1208 CGAATTGATTAATGGAAGTCAATGATTTACTAACAAAGAGGTTAAACAAGATGCTAT 1267
Db 5390 ---AATGATTAATGGAAGTCAATGATTTACTAACAAAGAGGTTAAACAAGATGCTAT 5334
QY 1268 A 1268
Db 5333 A 5333

RESULT
US-08-956-171B-142/C
; Sequence 142, Application US/08956171B
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gail H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
;
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
```

STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956.171B
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009.861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Hoover, Kenley K.
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 610-5790
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 142:
SEQUENCE CHARACTERISTICS:
LENGTH: 7588 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 142:
US-08-956-171B-142

Query Match 75.5%; Score 957.6; DB 30; Length 7588;
Best Local Similarity 93.1%; Pred. No. 1.6e-203;
Matches 1118; Conservative 0; Mismatches 62; Indels 21; Gaps 11;

QY 68 CAAAGTTTGGACATTACCAATATGATTCGGATGAGTCAATGNCANCGGTGTTAA 127
DB 6512 CCAAGGTTGGACATTAACAAATATGATTCGAT-AGTCAAAATGACAAACGGGTGTTAA 6454
QY 128 TAACTAGCAATGNTGNAATGATAGTAGNCAAGTGGCATACAGTTCCTCNAAT 187
DB 6453 TAACTAGCAATGNTGNAATGATAGTAGNCAAGTGGCATACAGTTCCTCNAAT 6394
QY 188 ATATGATCCTTAGAATGAGTGAACCAACAGCGCAGTTACATCAATAATTTGTAAC 247
DB 6393 ATATGATCCTTAGAATGAGTGAACCAACAGCGCAGTTACATCAATAATTTGTAAC 6334
QY 248 TAGAAGATTAATAGAGACCGCTATAGAGACGATTTGAAGTTGATTTAATGCTG 307
DB 6333 TAGAAGATTAATAGAGACCGCTATAGAGACGATTTGAAGTTGATTTAATGCTG 6274
QY 308 TTAGTAAGATCATATCATATGATGAGTGGCTATAGTACATGATTAATTAATAAACCG 367
DB 6273 TTAGTAAGATCATATCATATGATGAGTGGCTATAGTACATGATTAATTAATAAACCG 6214
QY 368 TCATTAAATGTTTTTTAGAAAACATATAGTATCAATTTAATAGTAGTACATCTACG 427
DB 6213 TCATTAAATGTTTTTTAGAAAACATATAGTATCAATTTAATAGTAGTACATCTACG 6154
QY 428 TACTCAAAATATCTATTAACAATTTTCAATATATATTTCTTGGGGCAGGTTGAATTCGA 487
DB 6153 TACTCAAAATATCTATTAACAATTTTCAATATATATTTCTTGGGGCAGGTTGAATTCGA 6094
QY 488 ACCGGCATTAATAAAGCTGGAGCTGCTATATGTTTCAATATAGTGGCGATCTAGT 547
DB 6093 ACCGGCATTAATAAAGCTGGAGCTGCTATATGTTTCAATATAGTGGCGATCTAGT 6034
QY 548 GAGATTCTAGAGCCGACGATTAAAGTCTGATGGAGAAAGATGTATATATGACAAA 607

DB 6033 GAGATTCTAGAGCCGACGATTAAAGTCTGATGGAGAAAAGATGTATATGACAAA 5974
QY 608 GATAATGAGCGTATTTGTAAATGTGTCAAAATAGCGTTATTAACGATTAATTTTC 667
DB 5973 GATAATGAGCGTATTTGTAAATGTGTCAAAATAGCGTTATTAACGATTAATTTTC 5914
QY 668 TCCTTCATCTTAATTCATCATGATGTGAGATTTTGTATAGAGGTGATCATTTGAGT 727
DB 5913 TCC-TTGATCTTAATTCATCATGATGTGAGATTTTGTATAGAGGTGATCATTTGAGT 5856
QY 728 CAATTATGATTAATGCGATTCGAATCCAAATATGTTACAGGTCAANACGTTGTTAT 787
DB 5855 CAATTATGATTAATGCGATTCGAATCCAAATATGTTACAGGTCAANACGTTGTTAT 5796
QY 788 CCACCGGTGGCGCTGTTAGTAAATGAAGGTGAGTGTGTTGTTGGTGGCACACTTG 847
DB 5795 CCACCGGTGGCGCTGTTAGTAAATGAAGGTGAGTGTGTTGTTGGTGGCACACTTG 5736
QY 848 AGAAAGGTGACACATGCGAGGTTCAAGCACTTGTATGCGACACAAANATGCTGAA 907
DB 5735 AGAAAGGTGACACATGCGAGGTTCAAGCACTTGTATGCGACACAAANATGCTGAA 5676
QY 908 GGTCCGACGATTTATATTAGTATGACCCATGATGATTTGTTGTTAACACACCTGT 967
DB 5675 GGTCCGACGATTTATATTAGTATGACCCATGATGATTTGTTGTTAACACACCTGT 5616
QY 968 GTTAAACAAATTTATGATTTGTTAGATAGCANAGTAGTATTAACACANAGCAATTC 1027
DB 5615 GTTAAACAAATTTATGATTTGTTAGATAGCANAGTAGTATTAACACANAGCAATTC 5556
QY 1028 CGTTAGACACATGCGTATGATGAGACGTTACGGGCTCCACGGTATTTGAGGGTTGAT 1087
DB 5557 CGTTAGACACATGCGTATGATGAGACGTTACGGGCTCCACGGTATTTGAGGGTTGAT 5499
QY 1088 GCGTTGATGATGAACGGCATTCACATTTATACCAAGACTTTTAAAGCAAAAAGCA 1147
DB 5498 GCGTTGATGATGAACGGCATTCACATTTATACCAAGACTTTTAAAGCAAAAAGCA 5448
QY 1148 AAGCACTTGGCACAAATTTACAGTGAAGTGTGAAGTTAGATGGGTAAACAAG 1207
DB 5447 AAGCACTTGGCACAAATTTACAGTGAAGTGTGAAGTTAGATGGGTAAACAAG 5391
QY 1208 CGAATGTAATAGGACAAAGTCAATGATTAACAAAGAGTTAAACAAGATGCTAT 1267
DB 5390 CGAATGTAATAGGACAAAGTCAATGATTAACAAAGAGTTAAACAAGATGCTAT 5334
QY 1268 A 1268
DB 5333 A 5333

RESULT 10
US-60-038-697-616
Sequence 616, Application US/60038697
GENERAL INFORMATION:
APPLICANT: Lagace, Robert E.
APPLICANT: Corley, Neil C.
APPLICANT: Russo, Frank D.
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF STAPHYLOCOCCUS AUREUS
TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USBS THEREOF
NUMBER OF SEQUENCES: 1027
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/038,697
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PO-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 616:
SEQUENCE CHARACTERISTICS:
LENGTH: 7003 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: SAC0616
US-60-038-697-616

Query Match      73.4% Score 931.6: DB 65: Length 7003:
Best Local Similarity 92.2% Pred. No. 1e-197:
Matches 1109: Conservative 0: Mismatches 72: Indels 22: Gaps 12:

QY 68 CAAGTTTGGACATTCACCAATATGATTCGATGAGGTCAATGNCANCGGTGTAA 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2138 CCAAGTTTGGACATTAACCAATATGATTCGAT -AGTCAATGACAAACGGTGTAA 2196

QY 128 TAAACTCGAAATGNTGTGAAATGATAGTAGANCAGTGGCGATACAGTCTCAATT 187
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2197 TAAACTCGAAATGNTGTGAAATGATAGTAGAGCAAGTGGCGATACAGTCTCAATT 2256

QY 188 ATATGATGCTTTAGATCGATGAGTGAACAACACGCGAGTTACATCAATATTTGTAA 247
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2257 ATATGATGCTTTAGATCGATGAGTGAACAACACGCGAGTTACATCAATATTTGTAA 2316

QY 248 TAGAAGATAATAAGAGACAGCTCTATAGAGACGAAATTTGAAGTTTAAATGTCG 307
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2317 TAGAAGATAATAAGAGACAGCTCTATAGAGACGAAATTTGAAGTTTAAATGTCG 2376

QY 308 TTATGATGATCATATATTAAGATGAGTGCCTATAGTACATCAATTTAATTAATTAAC 367
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2377 TTATGATGATCATATATTAAGATGAGTGCCTATAGTACATCAATTTAATTAATTAAC 2436

QY 368 TCATTAATGTTTTTTAGAAACATATAGTATCATTTTAAATGTAGTGCATACG 427
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2437 TCATTAATGTTTTTTAGAAACATATAGTATCATTTTAAATGTAGTGCATACG 2496

QY 428 TACTCAAAATATCTATTAACATTTTCAATATATATTTCTTGGGCGAGGTGAAAATGCCA 487
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2497 TACTCAAAATATCTATTAACATTTTCAATATATATTTCTTGGGCGAGGTGAAAATGCCA 2556

QY 488 ACCGGCAGTAAATTAACCCCTGCCAGCTGCTATATGTTTCAATATTAATGAGTACTAGT 547
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2557 ACCGGCAGTAAATTAACCCCTGCCAGCTGCTATATGTTTCAATATTAATGAGTACTAGT 2616

QY 548 GAGATTGTAGAGCGAGAGTTAAAGTGTGAGTGGAGAAAGAAATGTTAATTAATGAGACAAA 607
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2617 GAGATTGTAGAGCGAGAGTTAAAGTGTGAGTGGAGAAAGAAATGTTAATTAATGAGACAAA 2676

QY 608 GATAATGTACGATTTGTAAAAATGTGTACAAATAGGCTTATTTAA -CGATAAATTTT 666
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2677 GATAATGTACGATTTGTAAAAATGTGTACAAATAGGCTTATTTAA -CGATAAATTTT 2736

QY 667 CTCCTTTGCAT -CTTATTTCAATGATGAGGATTTTGTTTAATAGAGGATCATATTGA 725
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2737 CCCCCCTGCATCCTTAAATTCATGATGTGAGGA -TTTTTGTTTAATAGAGGATCATATTGA 2795

QY 726 GTCAATTTATGATTAATGCGATTCACTTCCAAATATGTATGATCAAGTCAANACAGGTGTTA 785

```

```

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 2796 GTCAATTTATGATTAATGCGATTCACTTGCAATATGTATGACAGGTCAANACAGGTGTTA 2855

QY 786 ATCCACCCCTGGCGCTGTTGTATGATTAATGAGTGAATGTTGTATGGTGCACACT 845
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 2856 ATCCACCCCTGGCGCTGTTGTATGATTAATGAGTGAATGTTGTATGGTGCACACT 2915

QY 846 TGAGAAAGGTGACAAAGCATGCGAGGTTCAACACTTGATATGACCAACAAATGCTG 905
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 2916 TGAGAAAGGTGACAAAGCATGCGAGGTTCAACACTTGATATGACCAACAAATGCTG 2975

QY 906 AAGTGGCAGATTTATATTACGTTAGAGCCATGTAGTCAATTTTGGTTCAACACACCT 965
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 2976 AAGTGGCAGATTTATATTACGTTAGAGCCATGTAGTCAATTTTGGTTCAACACACCT 3035

QY 966 GTGTTAACAAATATTGATTTGTAAGTATGACCAAAATAGTATTTACCAACAAAGCAAT 1025
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 3036 GTGTTAACAAATATTGATTTGTAAGTATGACCAAAATAGTATTTACCAACAAAGCAAT 3094

QY 1026 TCCGTTAGACACATGGGATGATGAGACGTTACGGGCTCCACGGTATTTGAGGGTTGAA 1085
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 3095 T-CGTTAGATACAT -GGTATGATACGTTACGGGATTCACGGTATTTGAGGGTTGAA 3152

QY 1086 TTGCGTTGATGATGAACGGGATTCACCAATTAATACCAAGCTTTTAAACCAAAAG 1145
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3153 TT-----GATGATGAACGTTCAATCAATTAATACCA--AGACTTTTAAAGCAAAAG 3203

QY 1146 CAAGCACTTGCACAAATTTACAGTGAAGTNTCTTGAAGTTTATGATGGTAAACAA 1205
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3204 CAAGCACTTGCACAAATTTACAGTGAAGTNTCTTGAAGTTTATGATGGTAAACAAAG 3260

QY 1206 AGCGAATGTAATGGAACAAAGTCAATGATGATTAATCAACAAAGGTTAAACAGATGCT 1265
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3261 CG---AATGATTAATGCAACAAAGTCAATGATGATTAATCAACAAAGGTTAAACAGATGCT 3317

QY 1266 ATA 1268
   |||
DB 3318 ATA 3320

RESULT 11
US-60-046-714-618
; Sequence 618, Application US/60046714
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert E.
; APPLICANT: Colley, Neil C.
; APPLICANT: Russo, Frank D.
; APPLICANT: Health, Joe D.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF STAPHYLOCOCCUS AUREUS
; TITLE OF INVENTIONS: GENOME, FRAGMENTS THEREOF, AND USES THEREOF
; NUMBER OF SEQUENCES: 1050
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/046,714
; FILING DATE: HERewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PM-0001-1P P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555

```

TELEFAX: (415) 845-4166
: INFORMATION FOR SEQ ID NO: 618:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7004 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: IMMEDIATE SOURCE:
: CLONE: SAUIC618
US-60-046-714-618

Query Match 73.4%; Score 931.6; DB 67; Length 7004;
Best Local Similarity 92.2%; Pred. No. 1e-197;
Matches 1109; Conservative 0; Mismatches 72; Indels 22; Gaps 12;

Oy 68 CAAGTTTGGAACTTACCAATATGATTCGATGAGTGCATTAATGNCANCGTGTAA 127
Db 2138 CCAAGTTTGGAACTTACCAATATGATTCGAT-AGGTCAATATGACAAAGCGTGTAA 2196
Oy 128 TAACTAGCAATGNTGTGNAATGATAGTANCAAGTGGCATACAGTTCNAT 187
Db 2197 TAACTAGCAATGNTGTGNAATGATAGTANCAAGTGGCATACAGTTCNAT 2256
Oy 188 ATATGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 247
Db 2257 ATATGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2316
Oy 248 TAGAAGATATTAAGAAGAAAGCTCTATAGAGAGCAATTAAGTGTATTAATGCTG 307
Db 2317 TAGAAGATATTAAGAAGAAAGCTCTATAGAGAGCAATTAAGTGTATTAATGCTG 2376
Oy 308 TTAGTAAGATATATCATATGATGATGATGATGATGATGATGATGATGATGATGAT 367
Db 2377 TTAGTAAGATATATCATATGATGATGATGATGATGATGATGATGATGATGATGAT 2436
Oy 368 TCATTAATGTTTATTTAGAAACATATATGATCATATTTAAATGATGATGATGATGAT 427
Db 2437 TCATTAATGTTTATTTAGAAACATATATGATCATATTTAAATGATGATGATGATGAT 2496
Oy 428 TACTCAATATATCTATACCAATTTCTATATATTTTGGGCGAGGCGTAATCCCA 487
Db 2497 TACTCAATATATCTATACCAATTTCTATATATTTTGGGCGAGGCGTAATCCCA 2556
Oy 488 ACCGCGAGTAATTAAGCTGCGACCTGCTATATGTTTCATATTAAGTGGCTGATCTAGT 547
Db 2557 ACCGCGAGTAATTAAGCTGCGACCTGCTATATGTTTCATATTAAGTGGCTGATCTAGT 2616
Oy 548 GAGATTTAGAGCCGACAGTTAAAGTCTGATGGAGAAAGATGTTAATTCGACAAA 607
Db 2617 GAGATTTAGAGCCGACAGTTAAAGTCTGATGGAGAAAGATGTTAATTCGACAAA 2676
Oy 608 GATAATGATGCTATTTGTAATAATGCTGTAATTTTA-CGATAAATTTTT 666
Db 2677 GATAATGATGCTATTTGTAATAATGCTGTAATTTTA-CGATAAATTTTT 2736
Oy 667 CTCCTTTCAT-CTTAATTCATGATGCTGAGTTCCTTTTCTTATAGAGTGCATTTGA 725
Db 2737 CTCCTTTCAT-CTTAATTCATGATGCTGAGTTCCTTTTCTTATAGAGTGCATTTGA 2795
Oy 726 GTCAATTTATGATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 785
Db 2796 GTCAATTTATGATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2855
Oy 786 ATCCACCCGTTGGCGCTGTTAGTTAATGAAGTAGAGATTGTTGTTGTTGTTGTTGTT 845
Db 2856 ATCCACCCGTTGGCGCTGTTAGTTAATGAAGTAGAGATTGTTGTTGTTGTTGTTGTT 2915
Oy 846 TGAGAAAAGTACAAAGATGCGGAGGTTCAAGCATTTGATATGCGACAAACAAATGCTG 905
Db 2916 TGAGAAAAGTACAAAGATGCGGAGGTTCAAGCATTTGATATGCGACAAACAAATGCTG 2975

Oy 906 AAGTGCGCAGCATTTATATATACGTTAGAGCCATGATGATTTTGGTTCAACACCACCT 965
Db 2976 AAGTGCGCAGCATTTATATATACGTTAGAGCCATGATGATTTTGGTTCAACACCACCT 3035
Oy 966 GTGTAAACAAATTTATTTATGTAAGATAGCANAAGTGTATTCNCNAACNAAGACAT 1025
Db 3036 GTGTAAACAAATTTATTTATGTAAGATAGCANAAGTGTATTCNCNAACNAAGACAT 3094
Oy 1026 TCCGTTAGACACATGCGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1085
Db 3095 T-CGTTATATACACT-AGTATGATGATGATGATGATGATGATGATGATGATGATGAT 3152
Oy 1086 TTGCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1145
Db 3153 TT-----GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3203
Oy 1146 CAAGCAACTTGGCCAAATTTACAGTAAAGTNTCTTGAAGTGTAGATGGTAAACAA 1205
Db 3204 CAAGCAACT--GGCACAATTTACAGTAAAGTATC-TGCAAGTGTAGATGTTAAACAAG 3260
Oy 1206 AGCGAATGATATGAGACAAAGTCAATGATTTACTTAACCAAGAGGTTAAACAAGTGTCT 1265
Db 3261 CG---AATGATATGAGACAAAGTCAATGATTTACTTAACCAAGAGGTTAAACAAGTGTCT 3317
Oy 1266 ATA 1268
Db 3318 ATA 3320

RESULT 12
US-09-611-529-2079
: Sequence 2079, Application US/09611529
: GENERAL INFORMATION:
: APPLICANT: George H. Shimer, Jr.
: APPLICANT: George H. Miller
: APPLICANT: Karen J. Shaw
: APPLICANT: Roderica S. Hare
: TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
: FILE REFERENCE: 1034/1C963US1
: CURRENT APPLICATION NUMBER: US/09/611, 529
: CURRENT FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/417, 811
: PRIOR FILING DATE: 1999-10-14
: PRIOR APPLICATION NUMBER: US 09/353, 718
: PRIOR FILING DATE: 1999-07-14
: PRIOR APPLICATION NUMBER: US 09/266, 557
: PRIOR FILING DATE: 1999-03-11
: PRIOR APPLICATION NUMBER: US 09/266, 556
: PRIOR FILING DATE: 1999-03-11
: PRIOR APPLICATION NUMBER: US 09/266, 555
: PRIOR FILING DATE: 1999-03-11
: PRIOR APPLICATION NUMBER: US 09/266, 542
: PRIOR FILING DATE: 1999-03-11
: PRIOR APPLICATION NUMBER: US 09/266, 541
: PRIOR FILING DATE: 1999-03-11
: PRIOR APPLICATION NUMBER: US 09/037, 934
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: US 09/036, 720
: PRIOR FILING DATE: 1998-03-06
: PRIOR APPLICATION NUMBER: US 09/036, 338
: PRIOR FILING DATE: 1998-03-06
: PRIOR APPLICATION NUMBER: US 09/036, 334
: PRIOR FILING DATE: 1998-03-06
: PRIOR APPLICATION NUMBER: US 09/036, 221
: PRIOR FILING DATE: 1998-03-06
: PRIOR APPLICATION NUMBER: US 09/036, 137
: PRIOR FILING DATE: 1998-03-06
: PRIOR APPLICATION NUMBER: US 09/036, 082
: PRIOR FILING DATE: 1998-03-06
: PRIOR APPLICATION NUMBER: US 09/036, 081
: PRIOR FILING DATE: 1998-03-06
: PRIOR APPLICATION NUMBER: US 09/036, 079
: PRIOR FILING DATE: 1998-03-06

```

; PRIOR APPLICATION NUMBER: US 09/035,913
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/035,744
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 08/827,356
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: US 08/831,156
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: US 60/014,477
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: US 60/016,743
; PRIOR FILING DATE: 1996-05-02
; PRIOR APPLICATION NUMBER: US 60/020,016
; PRIOR FILING DATE: 1996-06-14
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO: 2078
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-611-529-2079
```

```

Query Match      32.8%; Score 417.2; DB 57; Length 1119;
Best Local Similarity 89.7%; Pred. No. 4,8e-83;
Matches 560; Conservative 0; Mismatches 44; Indels 20; Gaps 10;
```

```

QY 645 GCTTATTACATAAATTTCTCTTGCATCTTAATTCATGATGATGATTTTGG 704
DB 1 gcttattacataaatttctcttgcatttcc-ttcgacattcaatgagtgagga-ttttgg 58
QY 705 TTTATAGAGGTATCATTTGAGTCAATTTATGATTTATGCGATTCACTCCAAATATGG 764
DB 59 ttatagaggtatcatcttgagtcattatgatatgatactgcaacttgcaaatatg 118
QY 765 TACAAGGTACAAAGGTGTTATTCACCCGTTGGCGCTGTTGATTGAATGAAGTAGGA 824
DB 119 tacaaggtacaaggtgtaattcacccgcttgcgctgttgagtaataagtagga 178
QY 825 TTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 884
DB 179 ttgttggttggttggttggttggttggttggttggttggttggttggttggttggt 238
QY 885 ATATGACACAAACAAATATGCGATTTATATGCGATTTATGCGATTTATGCGATTTATG 944
DB 239 atatgacacaaacaaatgctgaaagtgcgagcatattatagcttagagccatgtagtc 298
QY 945 ATTTGGTTCAACACCACCTGTGTTAACAAAATTATGATTGTAAGATAGCANAAGTAG 1004
DB 299 atttggttcaacaccacctgtgtaacaaaattatgattgtaagtagcnaaagtag 358
QY 1005 TATTACNCAACANAAGACATTCCTTATACACACATGGGTATGAGAGTTACGGGGCT 1064
DB 359 ta-taagcaacaanaagacaatt-cgttagacacat-9gtgtagagagcttaacgggtc 415
QY 1065 CCACGGTATTATGAGGTTCATATGCGATTTATGCGATTTATGCGATTTATGCGATTT 1124
DB 416 ccgcatatgaggttgaatgcgtt-----gagtagaaagggatcaacattatacca-- 467
QY 1125 GACTTTTTTAAAGCAAAAGCAAAAGCAAACTTGCCCAAAATTACAGTAAAGTNTCTTGG 1184
DB 468 -agactttttaagcaaaaagcaaaagcaact--gccacaattatcagtgaaagtatc-tg 523
QY 1185 AAGAGTTATATGCGTAAACAAAGCGAATGATATGACAAAGTCAATGATTTACTATAA 1244
DB 524 caagttatagatgtaacaagcg--aatgataatgacaagtaacatgattactaaca 580
QY 1245 AAGAGTTATATGCGTCTATA 1268
DB 581 aagaggttaacaagaatgcttata 604
```

```

RESULT 13
US-08-827-356-1376
```

```

; Sequence 1376, Application US/08827356
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Robert S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS RELATED
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 5574
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corporation
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033-0530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/827,356
; FILING DATE: 01-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/014,477
; FILING DATE: 01-APR-1996
; APPLICATION NUMBER: 60/016,743
; FILING DATE: 02-MAY-1996
; APPLICATION NUMBER: 60/020,016
; FILING DATE: 14-JUN-1996
; INFORMATION FOR SEQ ID NO: 1376:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1119 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1...1119
US-08-827-356-1376
```

```

Query Match      32.8%; Score 416.2; DB 23; Length 1119;
Best Local Similarity 89.7%; Pred. No. 8e-83;
Matches 559; Conservative 0; Mismatches 44; Indels 20; Gaps 10;
```

```

QY 646 CTTATTAAACGATAAATTTTCTCTTGCATCTTAATTCATGATGATGATTTTGGT 705
DB 1 cttattaaacgataaatttctcttgcatttcc-ttcgacattcaatgagtgagga-ttttgg 58
QY 706 TTTATAGAGGTATCATTTGAGTCAATTTATGATTTATGCGATTTATGCGATTTATG 765
DB 59 ttatagaggtatcatcttgagtcatttcc-ttcgacattcaatgagtgagga-ttttgg 118
QY 766 ACAAGGTCAACAGGTGTTAATTCACCCGTTGGCGCTGTTGATTGAATGAAGTAGGAT 825
DB 119 acaaggtcaaacaggtgtaattcacccgcttgccctgtttagttantgaagtaggatt 178
QY 826 TGTGATTTGTTGTCACACTTGAGAAAAGTGACAAAGCATGCGAGGTTCAAGCACTTGA 885
DB 179 tgtgatttggttgcacacttgagaaaagtgacaaagcatgcgaggttcaagcacttga 238
QY 886 TATGACACAAACAAATATGCGAAGTGCGACGATTTATGCGTAGAGCCATGATAGCA 945
DB 239 tatgacacaaacaaatgctgaaagtgcgacgatttatgacgttagagccatgtagca 298
QY 946 TTTGTTCAACACCACCTGTGTTAACAAAATTATGATTGTAAGATAGCANAAGTAGT 1005
```

```

|||||
Db 299 TTTTGGTTCAACACACCCTGTTGTTAAACAAATATTGATTGTAAGATAGCAAAAGTAGT 358
Qy 1006 ATTACNCAACANAAGACAAATTCCTGTAGACACACATGGGTATGAGACGTTACGGGCTC 1065
Db 359 A-TACGCAACAAAGCAAAAT-T-CGTAGACACACAT--GGTATGAGACGTTACGGGCTCA 415
Qy 1066 CACGGTATTTGAGGCTGAATTTGGTGTGATGAGACGGGCATCACAAATATACCAAG 1125
Db 416 CGGATTTAGGTTAAATGCGTT-----GATGATGAGACGGGCATCACAAATATACCA--- 466
Qy 1126 ACTTTTAAAAAGCAAAAGCAAAAGCACTTGCCACAAATTTACAGTGAAGTTCTTGA 1185
Db 467 AGACTTTTAAAGCAAAAGCAAAAGCAACT--GCCACAAATTTACAGTGAAGTATAC-TGC 523
Qy 1186 AAGTTAGATGGGTAACCAAGCGAATTTGATATGACAAAGTCAATGGATTACTACAA 1245
Db 524 AAGTTTATGATGTAACCAAGCG--AAATGATTAATGACAAAGTCAATGGATTACTACAA 580
Qy 1246 AGAGGTTAAACAGATGCTCTATA 1268
Db 581 AGAGGTTAAACAGATGCTCTATA 603

RESULT 14
PCT-US98-24857-1
: Sequence 1, Application PC/TUS9824857A
: GENERAL INFORMATION:
: APPLICANT: SmithKline Beecham Corporation
: TITLE OF INVENTION: Novel ribG
: FILE REFERENCE: P50444-9
: CURRENT APPLICATION NUMBER: PCT/US98/24857A
: CURRENT FILING DATE: 1998-11-23
: EARLIER APPLICATION NUMBER: 08/978,456
: EARLIER FILING DATE: 1997-11-25
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 1029
: TYPE: DNA
: ORGANISM: Staphylococcus aureus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)...(1026)
: PCT-US98-24857-1

Query Match 27.8%; Score 352.2; DB 1; Length 1029;
Best Local Similarity 88.4%; Pred. No. 1.6e-68;
Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;

Qy 734 ATGATTAATGCGATTCACACTCCAAATATGTTACAGGTGCANACAGGTGTTAATCCACC 793
Db 1 atgattatgcatcactcctgcataatggtacaaaggtcacaagtgtaataaccgcc 60
Qy 794 GTTGCGCGTGTAGTTAATGAAGTAGATGATGTTGGTATTTGGTGCACACTTGAGAAAA 853
Db 61 gtctggcgctgttagttaatgaagtagatgattgtgtatgtgtcaccacttgagaaaa 120
Qy 854 GGTGACAGCATCGGAGGTTCAGACACTTGATATGGCAACAACAAATGCTGAAGTGGC 913
Db 121 ggtgacaagcgtggaggttcacagcactgatatgcaacaacaatgctgaaggtggc 180
Qy 914 ACGATTAATATTACGTTAGACCATGTATGTCATTTGGTTCACACCACTGTGTTAAC 973
Db 181 acgattataatacgttagagccatgtgcatcttggttcaacacccctgtgttaac 240
Qy 974 AAAATTTATGTTTATAGATAGCANAACTAGTATTACNCAACANAAGCAATTCCTTAG 1033
Db 241 aaatatttgattgtaagatagcaaaagtagta-tacgcaacaagaacaaalt-cgttag 298
Qy 1034 ACACACATGGGTATGAGACGCTTACGGGGCTCCACGCTATTGAGGTTGATTTGCGTTG 1093
|||||
```

```

Db 299 acaacacat-gtgtatgagacgtctacggtctacagctatgaggttgatgctt----- 351
Qy 1094 GATGATGAAGGGGCATCACAAATATATACCAAGACTTTTAAAGCAAAAGCAAGCAAA 1153
Db 352 gatgatgaagggcatcacaaatatacacaagactttt---aaagcaaaagcaagcaaa 408
Qy 1154 CTTCGCCAAATTTACAGTGAAGTTCCTTGAAGTTTAGATGGTAAACAAAGCGAAT 1213
Db 409 ct-gccacaataatcagtgaaagtagtc-tgcaagtttagatgtataacaagcgt---aat 462
Qy 1214 GATPATGACAAAGTCAATGATTTACTACAAAGAGGTTTAAACAAAGATGCTCTATA 1268
Db 463 gataatgacaagaatgataatgtattacttaacaagaggtttaacaagatgtcata 517

RESULT 15
US-09-376-633-1
: Sequence 1, Application US/09376633
: GENERAL INFORMATION:
: APPLICANT: Palmer, Leslie M.
: TITLE OF INVENTION: Novel ribG
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechert Price & Rhoads
: STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
: CITY: Philadelphia
: STATE: PA
: COUNTRY: US
: ZIP: 19103
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/376,633
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/978,456
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Dickinson, Todd O
: REGISTRATION NUMBER: 28,354
: REFERENCE/DOCKET NUMBER: P50444-9
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 215-994-2252
: TELEFAX: 215-994-2222
: TELEX:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1029 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: US-09-376-633-1

Query Match 27.8%; Score 352.2; DB 103; Length 1029;
Best Local Similarity 88.4%; Pred. No. 1.6e-68;
Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;

Qy 734 ATGATTAATGCGATTCACACTCCAAATATGTTACAGGTGCANACAGGTGTTAATCCACC 793
Db 1 atgattatgcatcactcctgcataatggtacaaaggtcacaagtgtaataccacc 60
Qy 794 GTTGCGCGTGTAGTTAATGAAGTAGATGTTGGTATTTGGTGCACACTTGAGAAAA 853
Db 61 gtctggcgctgttagttaatgaagtagatgattgtgtatgtgtcaccacttgagaaaa 120
Qy 854 GGTGACAGCATCGGAGGTTCAGACACTTGATATGGCAACAACAAATGCTGAAGTGGC 913
Db 121 ggtgacaagcgtggaggttcacagcactgatatgcaacaacaatgctgaaggtggc 180
|||||
```



```

QY 914 ACGATTATATACGTTAGAGCCATGATGATTTTGGTCCACACCCCTGTGTTAC 973
    |||||||
Db 181 ACGATTATATACGTTAGAGCCATGATGATTTTGGTCCACACCCCTGTGTTAC 240
    |||||||
QY 974 AAAATTATGATTGTAGATAGCANNAGTAGTATTACNCAACANAGACAATTCCGTTAG 1033
    |||||||
Db 241 AAAATTATGATTGTAGATAGCANNAGTAGTATTACNCAACANAGACAATTCCGTTAG 298
    |||||||
QY 1034 ACACACATGGGTGATGAGACGTTACGGGGCTCCACGGTATTGAGGGTTGAATTGCGTTG 1093
    |||||||
Db 299 ACACACAT-GGTGATGAGACGTTACGGGGCTCCACGGTATTGAGGGTTGAATTGCGTTG 351
    |||||||
QY 1094 GATGATGAAAGGCGATCACAATTATACCAAGACTTTTAAAGCAAAAGCAAGCAAA 1153
    |||||||
Db 352 GATGATGAAAGGCGATCACAATTATACCAAGACTTTT--AAAGCAAAAGCAAGCAAA 408
    |||||||
QY 1154 CTGGCCACAATAATTACAGTGAAGTNTCTGAAGTTTAGTGGGTTAAACAAAGCGAATT 1213
    |||||||
Db 409 CT--GCCACAATAATTACAGTGAAGTATC-TGCAAGTTTAGTGGTAAACAAAGCG--AAT 462
    |||||||
QY 1214 GATATGGAACAAGTCAATGATTACTTAACAAGAGGTTAAACAGATGCTATA 1268
    |||||||
Db 463 GATAATGGAACAAGTCAATGATTACTTAACAAGAGGTTAAACAGATGCTATA 517
    |||||||

```

Search completed: November 19, 2000, 05:57:52
 Job time: 8499 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 00:50:17 ; Search time 2632.38 seconds
(without alignments)
1707.291 Million cell updates/sec

Title: US-08-978-456-1
Perfect score: 1029
Sequence: 1 ATGGATTATGCGATTCAACT.....TAACCTTAGCAAGAAGTGA 1029

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1033670 seqs, 2183789903 residues

Total number of hits satisfying chosen parameters: 2067340

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GeneBml:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl1:*
10: gb_pr2:*
11: gb_pr3:*
12: em_fun:*
13: em_hum1:*
14: em_hum2:*
15: em_in:*
16: em_om:*
17: em_or:*
18: em_ov:*
19: em_pat:*
20: em_ph:*
21: em_pl:*
22: em_ro:*
23: em_sts:*
24: em_sy:*
25: em_un:*
26: em_v1:*
27: gb_hcg1:*
28: gb_hcg2:*
29: gb_in1:*
30: gb_in2:*
31: em_ba1:*
32: em_ba2:*
33: em_hum3:*
34: em_hum4:*
35: gb_pr4:*
36: gb_hcg3:*
37: gb_hcg4:*
38: gb_hcg5:*
39: gb_hcg6:*
40: gb_hcg7:*
41: em_hcg1:*
42: em_hcg2:*
43: em_hcg3:*

44: em_hum5:*
45: gb_pl3:*
46: gb_pr5:*
47: gb_hcg8:*
48: gb_hcg9:*
49: gb_hcg10:*
50: gb_hcg11:*
51: gb_hcg12:*
52: gb_hcg13:*
53: gb_hcg14:*
54: gb_in3:*
55: gb_hcg15:*
56: gb_hcg16:*
57: gb_hcg17:*
58: em_hcg4:*
59: em_hcg5:*
60: em_hcg6:*
61: em_hcg7:*
62: em_hum6:*
63: gb_hcg18:*
64: gb_hcg19:*
65: gb_ba3:*
66: em_hcg8:*
67: em_hcg9:*
68: em_hcg10:*
69: em_hcg11:*
70: em_hcg12:*
71: em_hcg13:*
72: em_hcg14:*
73: em_hcg15:*
74: em_hcg16:*
75: em_hcg17:*
76: em_hcg18:*
77: em_hcg19:*
78: em_hcg20:*
79: em_hcg21:*
80: em_hcg22:*
81: em_hcg23:*
82: gb_pr6:*
83: gb_pr7:*
84: gb_hcg20:*
85: gb_hcg21:*
86: gb_hcg22:*
87: gb_hcg23:*
88: gb_ro:*
89: gb_sts1:*
90: gb_sts2:*
91: gb_sy:*
92: gb_un:*
93: gb_v11:*
94: gb_v12:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1029	100.0	3335	AR069579	AR069579-sequence--
2	516.8	50.2	3227	AF269646	AF269646 Staphyloc
3	516.8	50.2	3618	AF269345	AF269345 Staphyloc
4	340.6	33.1	2802	AF270130	AF270130 Staphyloc
5	155	15.1	3721	AF269926	AF269926 Staphyloc
6	120.4	11.7	6006	BSRIB	X51510 B.subtilis
7	120.4	11.7	28206	BACDIA	L09228 Bacillus su
8	120.4	11.3	218470	BSUB0013	299116 Bacillus su
9	116.4	11.3	14364	AE000675	AE000675 Aquifex a
10	112.2	10.9	4286	BARIBSENS	X95955 B.amycoliqu
11	95.2	9.3	18073	U32775	U32775 Haemophilus
12	91.4	8.9	11695	AE004298	AE004298 Vibrio ch

C	13	83.4	8.1	12811	1	AE001819	AE001819 Thermotog
	14	79	7.7	5851	65	AE009832	AE009832 Thermotog
	15	78.6	7.6	2849	2	ECNUSB	X64395 E. coli gene
	16	78.6	7.6	11692	1	AE000148	AE000148 Escherich
	17	78.6	7.6	139818	2	ECU82664	U82664 Escherichia
	18	77.4	7.5	105795	65	SYCPDNC	D64001 Synchocystis
	19	76.2	7.4	4312	1	APU27402	U27202 Actinobacill
	20	75.4	7.3	2759	1	AE002857	AE002857 Shigella
	21	67.8	6.6	11456	1	AE002277	AE002277 Chlamydia
	22	67.6	6.6	18759	1	AE002257	AE002257 Chlamydop
	23	67.6	6.6	20386	1	AE001667	AE001667 Chlamydia
	24	67.6	6.6	327167	1	AE002548	AE002548 Chlamydop
	25	66	6.4	11544	1	AE001343	AE001343 Chlamydia
	26	65.4	6.4	14376	1	AE002531	AE002531 Neisseria
	27	62.4	6.1	349601	2	NMA282491	AL162753 Neisseria
	28	60.4	5.9	2307	2	BHE133928	AL133928 Bartonell
	29	60.4	5.9	83779	45	SC6011856	AL11856 Saccharom
	30	60.2	5.9	86827	54	PEPMAL3P5	AL034556 Plasmidiot
	31	59.6	5.8	11404	1	AE003934	AE003934 Xylella f
	32	58	5.6	60604	49	AE023466	AE023466 Homo sapi
	33	56.4	5.5	150236	35	HS102111	AL021877 Human DNA
	34	56.4	5.5	196336	27	AC005506	AC005506 Plasmidiot
	35	55.6	5.4	7218	5	I66494	I66494 Sequence 14
	36	55.6	5.4	162160	38	AC011817	AC011817 Homo sapi
	37	54.6	5.4	177008	36	AC011175	AC011175 Homo sapi
	38	54.6	5.3	1686	19	E08995	E08995 DNA encodin
	39	54.4	5.3	53932	49	AC023371	AC023371 Homo sapi
	40	53.6	5.2	182827	27	AC005505	AC005505 Plasmidiot
	41	53.6	5.2	256172	27	AC005139	AC005139 Plasmidiot
	42	53.4	5.2	192581	87	PEMAL13P1	AL048180 plasmidiot
	43	51.8	5.0	53932	49	AC023371	AC023371 Homo sapi
	44	51.6	5.0	3804	4	YSCMTG604	Y56888 Saccharomy
	45	51.4	5.0	170143	53	AC045384	AC045384 Homo sapi

TS

RESULT	1								
LOCUS	AR069579								
DEFINITION	AR069579	3336 bp	DNA						18-FEB-2000
ACCESSION	AR069579	Sequence from patent US 5891672.							
VERSION	AR069579.1	GI:7220467							
KEYWORDS									
SOURCE	Unknown.								
ORGANISM	Unknown.								
REFERENCE	Unclassified.								
AUTHORS	1 (bases 1 to 3336)								
	Wang,M., Ward,J.M., Warren,R.,Lloyd, Nicholas,R.,Oakley,								
	Palmer,L.,Marie, Pratt,J.M., Knowles,D.,JustinCharles,								
	Lometto,M.,Arthur, Mooney,J., Black,M.Terence,								
	Burnham,M.,KarlRussell, Debouck,C.,Feldon,J.,Craig, Hodgson,J.E,								
	Jaworski,D.,Dee, Reichard,R.,Winfield, Rosenberg,M., Traini,C.M.,								
	and Zhong,Y.Y.1.								
TITLE	Polynucleotides encoding GMP cyclohydrogratase II (riBA)								
JOURNAL	Patent: US 5891672-A 7 06-APR-1999;								
FEATURES	Location/Qualifiers								
SOURCE	1..3336								
	/organism="unknown"								
BASE COUNT	1202 a	459 c	678 g	997 t					
ORIGIN									

Query Match	100.0%	Score 1029;	DB 5;	Length 3336;
Best Local Similarity	100.0%;	Pred. No. 1e-178;		
Matches 1029; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

20

Db	61	GTGGCGCTGTTGTAAGTTAATGAAGTACAGATTTGGTATTGGTGCACACTTCAGAAAA	120
Qy	121	GGTGACAGCATGGCGAGGTTCAAGCACTTGATATGACCAACAAAATGCTGAAGGTGC	180
Db	121	GGTGACAGCATGGCGAGGTTCAAGCACTTGATATGACCAACAAAATGCTGAAGGTGC	180
Qy	181	ACGATTATATATACGTTAGTAGGCATCTACTATTTTGGTTCAACACCACTCGTTTAAC	240
Db	181	ACGATTATATATACGTTAGTAGGCATCTACTATTTTGGTTCAACACCACTCGTTTAAC	240
Qy	241	AAAATTATGTTGTAAAGATAGCAAAAGTAGATACGCAACAAAAGCAATTCGTTAGAC	300
Db	241	AAAATTATGTTGTAAAGATAGCAAAAGTAGATACGCAACAAAAGCAATTCGTTAGAC	300
Qy	301	ACACATGGTGATGAGACGTTACGGGCTCACGGTATTAAGGTTGAATGCGTTGATGATGA	360
Db	301	ACACATGGTGATGAGACGTTACGGGCTCACGGTATTAAGGTTGAATGCGTTGATGATGA	360
Qy	361	CGGGCATCACAAATTATACCAAGACCTTTTAAAGCAAAAGCAACACGCGCAAAAT	420
Db	361	CGGGCATCACAAATTATACCAAGACCTTTTAAAGCAAAAGCAACACGCGCAAAAT	420
Qy	421	ACAGTGAAGATATCTGCACAACTTAGATGGTAAACAGCGAATGATATGACAAAGTCAA	480
Db	421	ACAGTGAAGATATCTGCACAACTTAGATGGTAAACAGCGAATGATATGACAAAGTCAA	480
Qy	481	TGGATTACTAACAAGAGGTTAAACAAGATGCTATTAAGTTAAACATCCAGACGACGA	540
Db	481	TGGATTACTAACAAGAGGTTAAACAAGATGCTATTAAGTTAAACATCCAGACGACGA	540
Qy	541	GTTGTAAGTGAAGACGTACAGTTGTAATAGATGATCCACAATACTACACGATTTCAA	600
Db	541	GTTGTAAGTGAAGACGTACAGTTGTAATAGATGATCCACAATACTACACGATTTCAA	600
Qy	601	GATGCAAAAAACCCATATAAAAAGTATATTTGCTAAAGTCTGGGAATATTCATTTTAACAG	660
Db	601	GATGCAAAAAACCCATATAAAAAGTATATTTGCTAAAGTCTGGGAATATTCATTTTAACAG	660
Qy	661	CAAAATTATCAAGTATGATCAACCAATTGGAATATACTGAAAATCCAAATTTACA	720
Db	661	CAAAATTATCAAGTATGATCAACCAATTGGAATATACTGAAAATCCAAATTTACA	720
Qy	721	AGCAATCAAAACACATATGAAATTAATTACTTGAAGTCTTGATTAACAAACATCTT	780
Db	721	AGCAATCAAAACACATATGAAATTAATTACTTGAAGTCTTGATTAACAAACATCTT	780
Qy	781	CACAATTATATATAAGAGAGGATGGACTTTCGTAAGTCCAGGAGGTCACCACTACT	840
Db	781	CACAATTATATATAAGAGAGGATGGACTTTCGTAAGTCCAGGAGGTCACCACTACT	840
Qy	841	TCAGAAATTCACATCTATTAATAGATGAATTAATTCCTATTAATGCCCGAATTAAT	900
Db	841	TCAGAAATTCACATCTATTAATAGATGAATTAATTCCTATTAATGCCCGAATTAAT	900
Qy	901	GGCGGATCTGGAAATTTATCAATTTATCAACCAATGATGATTTGAGATACCGAGATGC	960
Db	901	GGCGGATCTGGAAATTTATCAATTTATCAACCAATGATGATTTGAGATACCGAGATGC	960
Qy	961	AACCAATTGAATTTGTTCAATTCGAGATTATTAATCAAAATGTTAAATTAATTAATTCGA	1020
Db	961	AACCAATTGAATTTGTTCAATTCGAGATTATTAATCAAAATGTTAAATTAATTAATTCGA	1020

RESULT	2
AF269646/c	
LOCUS	AF269646 3227 bp DNA
DEFINITION	Staphylococcus epidermidis strain SRI clone step.1015d10 genomic sequence.

ACCESSION AF269646 GI:9623542
VERSION AF269646.1
KEYWORDS
SOURCE Staphylococcus epidermidis.
ORGANISM Staphylococcus epidermidis.
REFERENCE 1 (bases 1 to 3227)
AUTHORS Kimmerly, W.J., Taylor, J. David, Nelsen, A.J., Godlevski, M.M.,
Rubino, M.A., Nelson, F.J., Rivers, P.R., Torruella-Miller, I.,
Listebee, S., Ashanti, C., Altschuller, G., Mamou, L., Shepherd, N.S.,
Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and
Furdon, P.J.
TITLE Transposon-mediated sequencing of the Staphylococcus epidermidis
genome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3227)
AUTHORS Taylor, J. David, Kimmerly, W.J., Nelsen, A.J., Godlevski, M.M.,
Rubino, M.A., Nelson, F.J., Rivers, P.R., Torruella-Miller, I.,
Listebee, S., Ashanti, C., Altschuller, G., Mamou, L., Shepherd, N.S.,
Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and
Furdon, P.J.
TITLE Direct Submission
JOURNAL Submitted (22-May-2000) Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES
source location/Qualifiers
1..3227
/organism="Staphylococcus epidermidis"
/strain="SRI"
/db_xref="taxon:1282"
/clone="step.1015d10"

BASE COUNT 980 a 599 c 454 g 1194 t
ORIGIN
Query Match 50.2% Score 516.8; DB 1; Length 3227:
Best Local Similarity 69.5% Pred. No. 3.4e-85;
Matches 717; Conservative 0; Mismatches 312; Indels 3; Gaps 1;

QY 1 ANGATATGCGATTCACCTGCAAAATATGTCACAAAGTCAGGTGTTAATCCACC 60
DB 1918 ANGATATGCGATTCACCTGCAAAATATGTCACAAAGTCAGGTGTTAATCCACC 1859
QY 61 GTTGGCGCTGTGTACTTAATGAAGTAGATGTGTGTGTGTGTGTGTGTGTGTGT 120
DB 1858 GTTGGCGCTGTGTACTTAATGAAGTAGATGTGTGTGTGTGTGTGTGTGTGTGT 1799
QY 121 GGTGACAGCATGCGAGGTTCAGGACTGATGATGACACAAATGCTGAAGTGC 180
DB 1798 GGTGACAGCATGCGAGGTTCAGGACTGATGATGACACAAATGCTGAAGTGC 1739
QY 181 ACGATTTATATACGTAGAGCCATGATGATGATGATGATGATGATGATGATGAT 240
DB 1738 ACGATTTATATACGTAGAGCCATGATGATGATGATGATGATGATGATGATGAT 1679
QY 241 AAAATTTATGATGTAAGATGACAAAGTAGATGACAAAGTAGATGACAAAGTAG 300
DB 1678 AAAATTTATGATGTAAGATGACAAAGTAGATGACAAAGTAGATGACAAAGTAG 1619
QY 301 ACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 1618 ACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1559
QY 361 CGGCGATTCACATTAATACCAAGCTTTTAAAGCAAAAGCAAGCAAGCTGACCA 420
DB 1558 ACGGCGATTCACATTAATACCAAGCTTTTAAAGCAAAAGCAAGCAAGCTGACCA 1499
QY 421 ACGGCGATTCACATTAATACCAAGCTTTTAAAGCAAAAGCAAGCAAGCTGACCA 480
DB 1498 ACGGCGATTCACATTAATACCAAGCTTTTAAAGCAAAAGCAAGCTGACCA 1439
QY 481 TGGATTACACAAAGGTTAAAGATGCTATTAAGTAAAGATGACATGACAGCA 540

DB 1438 TGGATTACACAAAGGTTAAAGATGCTATTAAGTAAAGATGACATGACATGAC 1379
QY 541 GTTGTACAGCAAGAGCTACGATGATGATGATGATGATGATGATGATGATGAT 600
DB 1378 GTTGTACAGCAAGAGCTACGATGATGATGATGATGATGATGATGATGATGAT 1319
QY 601 GATGCAAAAACCTTAAAGTAAATATGATGATGATGATGATGATGATGATGATGAT 660
DB 1318 GATGCAAAAACCTTAAAGTAAATATGATGATGATGATGATGATGATGATGATGAT 1259
QY 661 CAAATTTATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 1258 CAAATTTATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1199
QY 721 ACGATCAACACATATTAATATTAATATTAATATTAATATTAATATTAATATTA 780
DB 1198 ACGATCAACACATATTAATATTAATATTAATATTAATATTAATATTAATATTA 1139
QY 781 CAAATTTATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 1138 CAAATTTATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1079
QY 841 TCGAATTTCTCCATCT--ATTATATGATGATGATGATGATGATGATGATGATGAT 897
DB 1078 TCGAATTTCTCCATCT--ATTATATGATGATGATGATGATGATGATGATGATGAT 1019
QY 898 ATTGGCGATCTGGAATATTAATTAATTAATTAATTAATTAATTAATTAATTA 957
DB 1018 ATTGGCGATCTGGAATATTAATTAATTAATTAATTAATTAATTAATTAATTA 959
QY 958 GCGAACAATTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017
DB 958 GCGAACAATTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 899
QY 1018 CGAAGAGCTGA 1029
DB 898 CGAAGAGCTGA 887
RESULT 3
AF269345/c 3618 bp DNA BCT 03-AUG-2000
LOCUS Staphylococcus epidermidis strain SRI clone step.1002d07 genomic
DEFINITION
ACCESSION AF269345
VERSION AF269345.1 GI:9664666
KEYWORDS
SOURCE Staphylococcus epidermidis.
ORGANISM Staphylococcus epidermidis.
REFERENCE 1 (bases 1 to 3618)
AUTHORS Kimmerly, W.J., Taylor, J. David, Nelsen, A.J., Godlevski, M.M.,
Rubino, M.A., Nelson, F.J., Rivers, P.R., Torruella-Miller, I.,
Listebee, S., Ashanti, C., Altschuller, G., Mamou, L., Shepherd, N.S.,
Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and
Furdon, P.J.
TITLE Transposon-mediated sequencing of the Staphylococcus epidermidis
genome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3618)
AUTHORS Taylor, J. David, Kimmerly, W.J., Nelsen, A.J., Godlevski, M.M.,
Rubino, M.A., Nelson, F.J., Rivers, P.R., Torruella-Miller, I.,
Listebee, S., Ashanti, C., Altschuller, G., Mamou, L., Shepherd, N.S.,
Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and
Furdon, P.J.
TITLE Direct Submission
JOURNAL Submitted (19-May-2000) Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES
source location/Qualifiers
1..3618

```

/organism="Staphylococcus epidermidis"
/strain="SRI"
/db_xref="taxon:1282"
/clone="step.1002d07"
BASE COUNT      1099 a      709 c      518 g      1292 t
ORIGIN

```

```

Query Match      50.2%  Score 516.8; DB 1; Length 3618;
Best Local Similarity 69.5%; Pred. No. 3.3e-85;
Matches 717; Conservative 0; Mismatches 312; Indels 3; Gaps 1;

```

```

OY 1 ATGGATTATGCGATTCACCTGCAATATGCTACAGCTCAACAGCTGTTATCCACC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2701 ATGGATGATGCTATTCACATGCAACAAATGTAATGACAAACAGCTTATCCACA 2642
OY 61 GTTGGCGCTGTGAGTAATGAAGAGTAGATTTGTTGTTGTCACACTTGAGAAA 120
    || || || || || || || || || || || || || || || || || || || || ||
Db 2641 GTAGGATCCGTTGTTGTTAAAAACGCTAGCATTTAGGTTAGTGTCACATTTAAAAAG 2582
OY 121 GGTGACAGCATCGCGAGCTCAACACCTTGATATGACACAAACAAATGCTGAAGTCCG 180
    || || || || || || || || || || || || || || || || || || || || ||
Db 2581 GGAGATAAACAATGCCAGAGTACAACTTTGAAATGGCAGGTTTAAATACCCAAAGTCT 2522
OY 181 AGAATTATATGCTGAGCATGTAGTATTTGGTTCACACACCCCTGTTTAC 240
    || || || || || || || || || || || || || || || || || || || || ||
Db 2521 ACCATATACGTTCTTATGAACTTGCACACACCTGTTTACACCCCTTGTGTCAT 2462
OY 241 AAAATTATGATTTGTAGATGACAAAGATAGTATACGACAAAGACAAATTCGTTAGC 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2461 AAAATCATTTGAAGCGGCGCATCTCTAAGSTCACTCTGTTTAAAGATACCTTTAGTA 2402
OY 301 ACACATGCTGATGAGACCTTACGGGCTACGGTATGAGCTTGAATGCGTGTGATGAA 360
    || || || || || || || || || || || || || || || || || || || || ||
Db 2401 AGTAGGGTGCAGGATTTCTAGAGAACCTGTATAGAGGTTGAATTCATATATATATA 2342
OY 361 CGGCGATCACAATTATACCAAGACTTTTAAAGCAAAAGCAACACTGCCACAATT 420
    || || || || || || || || || || || || || || || || || || || || ||
Db 2341 AATGACAGTCAATTTACCGGCTGCTTTTCTACTGTTAAAGCAAGAACTTCCAGAACTA 2282
OY 421 ACAGTGAAGATATCTGCAAGTTAGATGTAACAAAGCAAGCAATATGACAAAGTCAA 480
    || || || || || || || || || || || || || || || || || || || || ||
Db 2281 ACTGTAAGAGTCTCATCTAGTCTAGATGTAACAAAGCAAGCAACTTATATGAAGTAAG 2222
OY 481 TGGATTTACCAAGAGGTTAAACAAATGCTCTATAGTTAAAGCATCGACAGCAGCA 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2221 TGGATTAACAAACAAAGAGTTAAAGAAAGATGTTATCAATTAAGCATGACGATGATGCA 2162
OY 541 GTGTTAACTGGAAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGCA 600
    || || || || || || || || || || || || || || || || || || || || ||
Db 2161 GTTATTTACTGGCGCTAGAACCATTTGAACAGACATCCATTTGTATACACCAAGGTTCT 2102
OY 601 GATGAAAAAACCTTATTAAGTAATTTGCTAAGCTGGGAATATTCATTTTATATCAG 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2101 GATGAAAGCATCGATTCGAGTATATCTTCTTAAGAAAGGTCAACTGGATTTATATCAA 2042
OY 661 CAATTTATCAAGATGATCAACACCAATTTGATATATCTGATGAAATCCAAATTTACA 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2041 CAATATTTTAAAGTACTGACGAGATGATGATTTACCTGTAAGAAATGAATAATTAATA 1982
OY 721 AGCATCAACACATATGATGATTTACTTGAAGCTTGTGATTTAACAACAATTTCT 780
    || || || || || || || || || || || || || || || || || || || || ||
Db 1981 ACAATATTAATTTTATTAATTAATTAATTAATTAATTTGATGATGATGATGATGATTA 1922
OY 781 CACAATTTATTAAGAGAGGATTTGGAATTTGCTAGTGCAGGCGAGCTCAACCACTACT 840
    || || || || || || || || || || || || || || || || || || || || ||
Db 1921 CAAGACTATATCAAGAGGATTTGGGAACCTGCTAGTGCAGGCGAGCTCAACCACTACT 1862
OY 841 TCAGATTTCTCACTCT--ATTATATAGATGAATTTATCTCTATTTATGCCCGGAATTA 897
    || || || || || || || || || || || || || || || || || || || || ||
Db 1861 TCTCATATTTCTCCATCCAAACATCTTAATGAACATCTTTATATATATGATGATGATGAT 1802
OY 898 ATTGGCGATCTGGAATTTATCAACAAATGATGATGATGATGATGATGATGATGAT 957

```

```

Db 1801 ATTGGGTTCTGCGCAACATCATTTTATAGACTGACAGGCTGATTTGCTGAA 1742
OY 958 GCGAACCAATTTGAATTTGTTATTCGAGTATTTAATCAAAATGTAATTAATTTACTTTA 1017
    || || || || || || || || || || || || || || || || || || || || ||
Db 1741 GCAACTCAATTTGAATTTGTTATTCGAGTATTTAATCAAAATTTAATTAATTTAATTA 1682
OY 1018 CGAAGAAGATGA 1029
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1681 CGAAGAAGATGA 1670

```

```

RESULT 4
AF270130 2902 bp DNA BCT 01-AUG-2000
LOCUS
DEFINITION
Staphylococcus epidermidis strain SRI clone step.1051a10 genomic
sequence.
ACCESSION
AF270130
VERSION
AF270130.1 GI:9624037
SOURCE
Staphylococcus epidermidis.
ORGANISM
Staphylococcus epidermidis.
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Staphylococcus.

```

```

REFERENCE
1 (bases 1 to 2902)
Kimberly,W.J., Taylor,J.David, Nelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenbee,S., Ashanti,C., Altschuller,G., Mamou,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
Transposon-mediated sequencing of the staphylococcus epidermidis
genome
2 (bases 1 to 2902)
Taylor,J.David, Kimberly,W.J., Nelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenbee,S., Ashanti,C., Altschuller,G., Mamou,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
Direct Submission
Submitted (22-MAY-2000) Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
Drive, Research Triangle Park, North Carolina 27709-3398, USA
location/Qualifiers
1..2902
/organism="Staphylococcus epidermidis"
/strain="SRI"
/db_xref="taxon:1282"
/clone="step.1051a10"
BASE COUNT      1009 a      382 c      507 g      1004 t
ORIGIN

```

```

FEATURES
source
1..2902
/organism="Staphylococcus epidermidis"

```

```

Query Match      33.1%  Score 340.6; DB 1; Length 2902;
Best Local Similarity 69.0%; Pred. No. 5.1e-53;
Matches 466; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

```

```

OY 1 ATGGATTATGCGATTCACCTGCAATATGCTACAGCTCAACAGCTGTTATCCACC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2220 ATGGATGATGCTATTTCACTACGAAATATGTAATGCAAAACAGGTGTTATCCACCA 2279
OY 61 GTTGGCGCTGTGAGTAATGAAGAGTAGATTTGTTGTTGTCACACTTGAGAAA 120
    || || || || || || || || || || || || || || || || || || || || ||
Db 2280 GTAGGATCCGTTGTTTAAAAACGGTAGGATTTAGGTTAGTGACATTTAAAAAAG 2339
OY 121 GGTGACAGCATCGCGAGGTTCAAGCACTGATATGACACAAATGCTGAAGTCCG 180
    || || || || || || || || || || || || || || || || || || || || ||
Db 2340 GGAGATAAACAATGCCAGAGTACAAAGTATTTGAATGCGACAGTTAATACCAAGTGCT 2399
OY 181 AGAATTATATTAAGTATGAGCATGATGATTTGTTGTTCAACACACCTGTGTATAC 240
    || || || || || || || || || || || || || || || || || || || || ||
Db 2400 ACCATATACGTTTCATTAAGAACCTTGACACACCATGATGTTCAACACACCTGTGTGAT 2459

```

OY	241	AAAAATATGATTGTAAGATTACCAAAAGTAGTATCGCAACAAAGACAAATTGCTTAGAC	300
Db	2460	AAAAATCATTTGAAGCGGGCATATCTAAGGTCATCTATGCTTTAAAGATACTACTTTAGTA	2519
OY	301	ACACATGGTGATGAGACGTTTACGGGCTCAGGATTTAGAGTTGAATGCGCTTGATGATGA	360
Db	2520	AGTAAGGGTGACGAGATTCTGTAGAGAGAGCGTGGTATAGAGTTGAAATTCATATATATGA	2579
OY	361	CGGGCATCAACAATTATPACCAAGACTTTTTTTAAAGCAAAAGCAAGCACTGCCACAATT	420
Db	2580	AATGAGAGCTGCATTTATACCGTACTCTTTTCTACTGCTTAAAGAAACGAAATGCCAGAA	2639
OY	421	ACAGTGAAGATGATCGCAAGTTTAGTGGTAAACAAGCAACGATATGATGACAAAGTCA	480
Db	2640	ACTGTAAAGGTTCTACTCTTAGTGTAGTGTAAACAAGCAAGCAAGACTTTATGAAAGTA	2699
OY	481	TGGATTACTAACAAAGAGGTTAAACAAGATGCTCTATAGTTAAAGACATGACAGCAGCA	540
Db	2700	TGGATTACCAACAAAGAAAGTTAAAGAAAGATGTTTATTCATTTAAAGACATGAGCAT	2759
OY	541	GGGTAACTGGAAGACGTACACTTGATTAATAGATGCCAATATATCTACAGTATCA	600
Db	2760	GTTATTACTGGCGGTAGAACCAATTGAGGAGCAATCCATTTGATCAACACAGGGTCT	2819
OY	601	GATGCAAAAAACCTCATTAATAAATAATATTTGTCTAGTGGGAATTTATTTAATACG	660
Db	2820	GATGAAAGACATCCGATTCGATTTATCTCTTTCTTAAAGAAAGTCACGATTTTAATCA	2879
OY	661	CAAAATTTATCAAGAT	675
Db	2880	CAAAATTTTAAAGAT	2894
RESULT	5		
AF269926/c			
LOCUS	AF269926	3721 bp	DNA
DEFINITION	Staphylococcus epidermidis strain SRI clone step.1032802 genomic	sequence.	sequence.
ACCESSION	AF269926		
VERSION	AF269926.1	GI:9623827	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
BASE COUNT			
ORIGIN			

[illegible]

[illegible]

Query Match	11.7%	Score 120.4	DB 2	Length 6006
Best Local Similarity 51.0%				
Matches 317	Conservative 0	Mismatches 256	Indels 9	Gaps 1
BASE COUNT	1891 a	1167 c	1401 g	1547 t
ORIGIN				
Query Match				
1	ATGATTTATGCGATTCAACCTTGCAAAATATGGTACAGGTCAACAGGTGTTAATCCACC	60		
Db	1227	ATGAAGCTGGCGCTTGATCTTGCGAAGCGGCGGAGAGACAGCAATCCATCCGCTC	1286	
Qy	61	GTTGGCGCTGTTGTTAGTAAGACGTGGAATGTTGGTATTGGTGCACTGTGAGAAA	120	
Db	1287	GTCGGCGCTGTTGTTGTAAGAGCGGACAAATTCGGAATGGCGCCATTTAAATAT	1346	
Qy	121	GGTGCAAGCATGCGGAGGTTCAAGCACTTATATGAGCACAACAAATGTGAAAGTGGC	180	
Db	1347	GGTGAAGCTCAGCAGAAAGTTTATCCATTCATTCGCTGGAGCAGCATGAGAGGTC	1406	
Qy	181	ACGATTTATATTACGTAGAGCATGTAGTCATTTGGTTCAACACACACCTGTGTAC	240	
Db	1407	GACATTTACGTTACCTCGAAACCGGCGACCATTAACGGAACACCCCATGTGACAGA	1466	
Qy	241	AAATTTATGATGTAATGAATAGCAAAATAGATACGCAACAAAAG-----CAAT	291	
Db	1467	TTGATTTATCACTCTGTTATCAAAAGAGTGTTCGTGGCGATGAGAGATCTTAATCCGCTT	1526	
Qy	292	TGCTTAGACACACATGATGATGAGACGTTACGGGCTCAGGTTAGAGTTGAATGCGCTT	351	
Db	1527	GTCGCTGGAAGAGGATCAGCATGATGAAGAGCTGGCATTTAGSATAAGGAGGCATC	1586	
Qy	352	GATGATGAAGCGGCATCAATTTATACCAAGACTTTTAAAGCAAAAACCAAGCACTG	411	
Db	1587	CTGGCAGACACAGCGGAGAGGCTGTAATGAAAAATTTCTGCACCTTATAGAGCAGGCTT	1646	
Qy	412	CCACAAATTTACAGTAAAGTATCTGCAGTTTAAAGTGTAAACAGCAATGATTAATGA	471	
Db	1647	CGTACGTCACCTTAAAGCGGCTGCCAGCCTTGACGGCAAGATATGTCACAGCAGGGGT	1706	
Qy	472	CAAACTCATAGGATTAAGTAAAGAGGTTTAAACAGATGCTATTAAGTAAAGCATGGA	531	
Db	1707	GACAGCAATATGATACCTCAGAGGCTCCAGACAGATGCTCAGCATATACAGGAANCA	1766	
Qy	532	CACGACGAGAGTGTAACTGGAAGACGTACAGTTGATTTAGATGATCCAAATATATCTACA	591	
Db	1767	CACCAAGAGATTTTGTGGAGTTGGCACAGTGAAGCGGCAACATCCAGGTTAACTGCG	1826	
Qy	592	CGTATTCAAGATGAAAAAAC	613	
Db	1827	AGACTGCGGAATGTAAACAAAC	1848	
RESULT 7				
BACDIA	28206 bp	DNA	ECT	26-MAY-1995
LOCUS	Bacillus subtilis spova to sera region.			
DEFINITION	L09228			
ACCESSION	L09228.1	GI:410114		
VERSION				
KEYWORDS	3-dehydroxynate dehydratase; aroc gene; diaminoimelate decarboxylase; lysa gene; penicillin-binding protein; peptidyl-prolyl isomerase; phosphoglycerate dehydrogenase; ppib gene; response regulator; response regulator kinase; riba gene; ribb gene; ribg gene; ribh gene; ribt gene; riboflavin biosynthesis operon; sera gene; signal peptidase; slps gene; spoa gene; spovaf gene.			

SOURCE Bacillus subtilis (strain 168, sub-species Marburg) DNA.
ORGANISM Bacillus subtilis
Bacteria: Firmicutes: Bacillus/Clostridium group:
Bacillus/Staphylococcus group; Bacillus.
REFERENCE 1 (bases 1 to 28206)
AUTHORS Yamamoto,J., Shimizu,M. and Yamane,K.
TITLE Molecular cloning and analysis of nucleotide sequence of the
Bacillus subtilis lysA gene region using B. subtilis phage vectors
and a multi-copy plasmid. PubMed
JOURNAL Agric. Biol. Chem. 55 (6), 1615-1626 (1991)
MEDLINE 91345841
REFERENCE 2 (bases 1 to 28206)
AUTHORS Buchanan,C.E. and Ling,M.L.
TITLE Isolation and sequence analysis of dacB, which encodes a
sporulation-specific penicillin-binding protein in Bacillus
subtilis
JOURNAL J. Bacteriol. 174 (6), 1717-1725 (1992)
MEDLINE 92193254
REFERENCE 3 (bases 1 to 28206)
AUTHORS Sorokin,A., Zunsteln,E., Azevedo,V., Ehrlich,S.D. and Serrro,P.
TITLE The organization of the Bacillus subtilis 168 chromosome region
between the spoVA and serA genetic loci, based on sequence data
JOURNAL Mol. Microbiol. 10 (2), 385-395 (1993)
MEDLINE 95020538
FEATURES
source location/Qualifiers
1. .28206
/organism="Bacillus subtilis"
/strain="168"
/sub_species="Marburg"
/db_xref="taxon:1423"
1. .1239
/gene="spovAF"
<1. .1239
/gene="spovAF"
/codon_start=1
/transl_table=11
/protein_id="AA67472.1"
/db_xref="GI:410115"
/translation="VDIVENRLMAQVEKVTIDETTDVLSGLAVIVYEGAFATII
DRASTYGRNPEEDPTKEVYRGARGGVENVVATALLRRIRDERLRVAKTKGKESK
TDLSTICYIEDIADPDLVEYKEIASIDVGLTMAKVAEEFVNSVNPFLVRYTE
RDPVANNHVLGHHVILIVDTSPEVITPTTLFHHVQHAEEYRQAPSVGTFLRVRFEG
IASTFLFLPWFLEAFDLDPNMKFGINKPTHHILLOIFLADLGIEFLMAAHS
TPTALSTAGLAAVLIGQIAIEVGLFSEVILVYLSAIGFTTPTSEVLSLATNPS
CHRDTRCPISYKRAARRRLKSANMAASINSLOPIYMLPIFGKALMQVLYRTAKPG
AKVRPSIVHKNRLRQPTNS"
1158. .1164
/gene="spovAF"
/citation=[1]
1245. .1250
/gene="lysA"
/note="putative"
1245. .2664
/gene="lysA"
1266. .1272
/gene="lysA"
/note="putative"
1345. .2664
/gene="lysA"
/codon_start=1
/transl_table=11
/product="diaminopimelate decarboxylase"
/protein_id="AA67473.1"
/db_xref="GI:410117"
/translation="MFLGTSRONQHGLEIGVDALYLAEKYGPLVYDVALLER
AKSFOAFISAGLKAQVAAASKAFSSVAMQLAEEGLSDVYSGGELYTAVAAGPA
ERIHFGNNKSRLELMALREHIGCIYVDNFIALLDLCKETGSHSIDVLLITPGV
EATHDYITTGQDSKFGDLNNGOTERAIEOVLOJSEHIOILGVCHISLOJDTGFG
VLAEEKIFKKLDWRDSYSPVSKVNLGGFCGRYEDDEPILHATYVKKITEAVEN
ASRTGFDIPETWIEPGRSLVGDAQTLLYTGSKQEPVGRQYAVVDGGMNDIRPALY
QAKYEAAANRIGEAHDKTYSINGKCESGMDIMPIDLPEVKGDLLVFCGAYGY
SMANNYRIPRPVAVFVENGEAHLVVKRETYEDIVLDELPEFKYGVKQ"
1610. .1616

/gene="lysA"
/citation=[1]
1751. .1757
/gene="lysA"
/citation=[1]
2496. .2502
/gene="lysA"
/citation=[1]
2504. .2510
/gene="lysA"
/citation=[1]
2666. .2694
/note="ORFX19"
complement(2693. .3565)
/note="ORFX19"
/codon_start=1
/transl_table=11
/protein_id="AA67474.1"
/db_xref="GI:410118"
/translation="MKTIMGWLAAVLLIWWPKVSLADAAGDYVTLLGADLSFSDK
QKVLDEMNVPDNATVYVTNKEHEHYLGKYSISNAIGSRALSSSITAKGSGLTNE
THNIGITDEMILNALMTAGVDAAVYVTAPEVSGTALGELIAVEVSSDEAISD
VKQVANOELVYTSLEDKRGNEANAALIAKKEEPKAGGVPNKADIEKQVDDAASDL
NVTLDQSKNOQLVSLFNNKKNADIDMGQVSDLDKAKDKITKFIESDEGKNFIQKVID
FVFSIMNAIVSIFK"
3347. .3353
/citation=[1]
3413. .3419
/citation=[1]
3428. .3434
/citation=[1]
3503. .3509
/citation=[1]
3819. .4290
/gene="ppIB"
3819. .4250
/gene="ppIB"
/codon_start=1
/transl_table=11
/product="peptidyl-prolyl isomerase"
/protein_id="AA67475.1"
/db_xref="GI:410119"
/translation="MKTGFLELDGKNKIEFELYPPAAGTVAEFKLANEGFYDLTF
HVRIGFVSGGCPHGTGGPGYTIKCTEGNPTTHGALISMNAHAKDKGGSGFFI
VHEPQPHANGVTVGKVTSGLEFAKNNNGDVMEVVEG"
3982. .3988
/gene="ppIB"
/citation=[1]
4261. .4290
/gene="ppIB"
4287. .4320
/note="inverted repeat"
5264. .5467
/note="ORFX1"
/codon_start=1
/transl_table=11
/protein_id="AA67476.1"
/db_xref="GI:410120"
/translation="MQVLTGLLEQYGVAVLFVLMLELLALLGPEFGVLFVDEKLN
WVLSMLAAGLGCGSEHYVLDLV"
5445. .5834
/note="ORFX2"
/codon_start=1
/transl_table=11
/protein_id="AA67477.1"
/db_xref="GI:410121"
/translation="MSWIGYVGLGPAALOKAFASRGVLLAVLVILFFWTVRGDRKI
LESLFPLVIGLIEEGIRLRFHSLHVNAPLESHVLYVPPSBQTLAAVYIFGRCAV
LFSRSHKTRLOTVAFFCLLSYSSSG"
6022. .6027
/gene="sips"
/note="putative"

conflict

-35_signal

```

gene      6022.. 7323
/gene="sips"
-10_signal 6045.. 6051
/gene="sips"
-35_signal 6683.. 6688
/note="putative"
/note="sips"
-10_signal 6705.. 6710
/note="putative"
/note="sips"
CDS       6740.. 7294
/note="putative"
/gene="sips"
/transl_table=11
/product="signal peptidase"
/protein_id="AA67478.1"
/db_xref="GI:410122"
/translation="MKSENVSKKKSLLEMAKAIYIAVAVLALLINFIAPYVVDGDSM
YPLAHNRERVNNTVKYIGFEDRGDIYLVNGDDVHYKRIIGLPGDTVEKKNDOLYI
NGKRVDEPYLANKKRAKODGFHLLTDFGIVKVPDKNYFVGMGNRRNSDSRNGSLGL
FTKROIAGTSKFVFYFENEMRKIN"

terminator
-35_signal 7388.. 7393
/note="sips"
-10_signal 7410.. 7416
/note="putative; OREX4"
/note="putative; OREX4"
CDS       7529.. 7873
/note="OREX4"
/codon_start=1
/transl_table=11
/protein_id="AA67479.1"
/db_xref="GI:410123"
/translation="MGRIKTKITLLVLLLLLAGCYWYINDIEKDPVTAIGTSLSE
EEEYIOEYKVKTKIDGSEYHGVANGNRIIFNGKRLMODLSIDIEGDKIKAYFESKSR
IDGLIKAVKND"
terminator 7872.. 7911
/note="OREX4"
-35_signal 7937.. 7942
-10_signal 7960.. 7965
protein_bind 8126.. 8137
/note="putative"

Query Match      11.7%; Score 120.4; DB 1: Length 28206;
Best Local Similarity 51.0%; Pred. No. 6.3e-13;
Matches 317; Conservative 0; Mismatches 296; Indels 9; Gaps 1;

```

```

Db      8641 CTGCAGACGCGGAGAGAGGCTGAATGAATAAATTTCTGCACCTTATAGGACAGCGCTT 8700
Qy      412 CCACAATTACAGTGAAGAAAGTCTGCAAGTTAGATGCTAAACAGCGAATGATATAGA 471
Db      8701 CCGTACGTCAGCGTAAAGCGCGTCCAGCGCTTGACGGCAAGATAGTACCAGCAGCGGT 8760
Qy      472 CAAAGTCATGATTTACTTAACAAAGAGTAAACAGATGCTCTAAGTTAAGACATCGA 531
Db      8761 GACAGCAATGATCAGCTCAGAGGCTGCAAGACAGATGCTCGCAATACAGAAACAA 8820
Qy      532 CACGACGAGTGTCTACTGGAAGACGTACAGTTGAATTAAGATCCACATATACTACA 591
Db      8821 CACCAAGACATTTTATGTCGAGGTGGCCAGATGAAACCCGCAATCCGAGCTTACCTGC 8880
Qy      592 CGATTTCAGATGCAAAAACC 613
Db      8881 AGACTGCCGCAATGTACAAAAC 8902

RESULT      8
BSUB0013/c 218470 bp DNA BCT 26-NOV-1997
LOCUS      Bacillus subtilis complete genome (section 13 of 21): from 2395261
DEFINITION to 2613730.
ACCESSION 299116 AL009126
VERSION    299116.1 GI:2634723
KEYWORDS
SOURCE
ORGANISM   Bacillus subtilis.
REFERENCE 1 (bases 1 to 218470)
AUTHORS   Kunst,F., Ogasawara,N., Moszer,I., Albertini,A.M., Alloni,G.,
Aaevedo,V., Bertero,M.G., Bessières,P., Bolotin,A., Borchert,S.,
Boris,R., Boursier,L., Brans,A., Brann,M., Brignell,S.C.,
Bron,S., Brouillet,S., Brusch,C.V., Caldwell,B., Capuano,V.,
Carter,N.M., Choi,S.K., Codani,J.J., Conerton,I.F., Cummings,N.J.,
Daniel,R.A., Denizot,F., Devine,K.M., Dusterhoft,A., Ehrlich,S.D.,
Emmerson,P.T., Entlan,K.D., Errington,J., Fabret,C., Ferreri,E.,
Fouger,D., Fritz,C., Fujita,M., Fujita,Y., Fuma,S., Galizli,A.,
Galleron,N., Ghim,S.Y., Glaeser,P., Goffeau,A., Gollightly,E.J.,
Grandi,G., Guisepi,G., Guy,B.J., Haga,K., Halech,J., Harwood,C.R.,
Henaut,A., Hilbert,H., Holsappel,S., Hosono,S., Hullo,M.F.,
Itaya,M., Jones,L., Joris,B., Karamata,D., Kasahara,Y.,
Klaerr-Blanchard,M., Klein,C., Kobayashi,Y., Koetter,P.,
Koningstein,G., Krogh,S., Kumano,M., Kurita,K., Lapius,A.,
Lardinois,S., Lauber,J., Lazarevic,V., Lee,S.M., Levine,A., Liu,H.,
Masuda,S., Mauel,C., Medigue,C., Medina,N., Mellado,R.P.,
Mizuno,M., Moestl,D., Nakai,S., Noback,M., Noone,D., O'Reilly,M.,
Ogawa,K., Ogiwara,A., Oudega,B., Park,S.H., Paro,V., Pohl,T.M.,
Portetelle,D., Potvollik,S., Prescott,A.M., Presecan,E., Fujic,P.,
Purnelle,B., Rapoport,G., Rey,M., Reynolds,S., Rieger,M.,
Rivolta,C., Roche,E., Roche,B., Rose,M., Sadale,Y., Sato,T.,
Scanlan,E., Schleich,S., Schroeter,R., Scoffone,F., Sekiguchi,J.,
Sekowska,A., Seror,S.J., Serron,P., Shin,B.S., Solido,B.,
Sorokin,A., Tacconi,E., Takagi,T., Takahashi,H., Takemaru,K.,
Takeuchi,M., Yamakoshi,A., Tanaka,T., Tepstra,P., Tognoni,K.,
Tosato,V., Uchiyama,S., Vandenbol,M., Vanlier,F., Vassiroli,A.,
Viari,A., Wambolt,R., Wedler,E., Wedler,H., Weltetoger,T.,
Winters,P., Wipalt,A., Yamamoto,H., Yamane,K., Yasumoto,K., Yata,K.,
Yoshida,K., Yoshikawa,H.F., Zumbstein,E., Yoshikawa,H. and
Danchin,A.
TITLE      The complete genome sequence of the gram-positive bacterium
Bacillus subtilis
Bacillus subtilis
Nature 390 (6657), 249-256 (1997)
98044033
2 (bases 1 to 218470)
JOURNAL    Kunst,F., Ogasawara,N., Yoshikawa,H. and Danchin,A.
MEDLINE    Direct Submission
AUTHORS     Submitted (18-NOV-1997) I. Moszer, A. Danchin, Institut Pasteur,
REFERENCE   Regulation de l'Expression Genetique, 28 rue du Docteur Roux, 75724
JOURNAL     Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr,

```

adanching@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45 68 89 48

FEATURES

source

Location/Qualifiers
1. 218470

/organism="Bacillus subtilis"

/strain="168"

/db_xref="taxon:1423"

complement(7. 18)

terminator

/gene="cmk"

complement(7. 696)

gene

/gene="cmk"

complement(22. 696)

CDS

/gene="cmk"

/cc-number="2,7 4,14"

/note="alternate gene name: jofc, ypfC"

/codon_start=1

/transl_table=1

/product="cytidylate kinase"

/protein_id="CAB14221.1"

/db_xref="GI:2634724"

/db_xref="SWISS-PROT:P38493"

/translation="MEKRLSIAIDGPAAGKSTVAKIVAEKKSYIYIDGAMTRATY
ALQENVLDLDEKELAEKLRDIELITTDGKVFVNGDVEAIRTDEISNQVSLA
AKHRSVSEKVKRQOQGEKGVMDGRDITVLPAAEVKIFLASVEERAKRYEE
NVKGEVNYETLIEEIAARDKLDSEREVSPLRKAEDALEIDTSLSIQEVADKILEA
VEOKSR"

gene

complement(775. 951)

CDS

/gene="ypfB"

complement(775. 951)

/gene="ypfB"

/function="unknown"

/note="alternate gene name: jofB"

/codon_start=1

/transl_table=1

/protein_id="CAB14222.1"

/db_xref="GI:2634725"

/db_xref="SWISS-PROT:P38492"

/translation="MKTFERLLKILFLIQAILLLGVQLFHFQHIIEPYVSKVIQYEGV
DKMEENRRTFFH"

gene

complement(996. 1649)

CDS

/gene="ypfA"

complement(996. 1649)

/gene="ypfA"

/function="unknown"

/note="alternate gene name: jofA"

/codon_start=1

/transl_table=1

/protein_id="CAB14223.1"

/db_xref="GI:2634726"

/db_xref="SWISS-PROT:P38491"

/translation="MTEIGENVLLLEVIEEELIKKASKAVSINNELLIVPVVYTG
RTVILHNDMEVTVFVNGKDEVPYFRFISRTIGKVKDKIOMCLEPPPEKKRRIORQY
VRIDVAVDVOIQGNEEIRTLSTINISAGIYAVLADGLSFQSGESLRLIIRPEEHH
TRIDETEAIVVRIRFNDPSEKSRKMTLEYSEIAQDOALLQYCIROILNKRRAHE"

gene

complement(1711. 3094)

terminator

/gene="ypdB"

complement(1711. 1733)

CDS

/gene="ypdB"

complement(1742. 3094)

/gene="ypdB"

/function="unknown"

/note="alternate gene name: jofB, yzua

similar to hypothetical proteins"

/codon_start=1

/transl_table=1

/protein_id="CAB14224.1"

/db_xref="GI:2634727"

/db_xref="SWISS-PROT:P38490"

/translation="MIRGILAVIGISIVGTGYGWGYKEHOEKDAVILLHAENNYORAFH
ELTYQVDQLHDKIGTITLANSQKSLSPALIDWKRITSEAHNSYSQLPITLPNKKEE
LSKIGDGYSTKTSVRDLDPKPLKNETSTLNKLQOSEDIONELRHVHOLVPSKRLNW
MDVEMALASDEKQSDNTIINSFTVEKNVAFSTGDLGFSFTSKKEEGFSLKGG
OISEQEAQIAERFAPDDNYSIKVYKSGKKTNRDYSISMKDPDKAVIYMDITKKGK

gene

complement(3129. 4046)

CDS

/gene="slkB"

complement(3129. 4046)

/gene="slkB"

terminator

/note="alternate gene name: ypea"

/codon_start=1

/transl_table=1

/product="spore cortex lytic enzyme"

/protein_id="CAB14225.1"

/db_xref="GI:2634728"

/db_xref="SWISS-PROT:P50739"

/translation="KSKSGSIACILLESFTITPTINTETISAFSNQVIRQATGDDV
VELQARLQYNGYNGKIDGVYGMGIWAVRNQDQDGLVEVGLGAKKQQLICKSK
YREYVMEQLNGNFTHYGKIDPKYQTPSKAAQOKAQOAEARQKOPAEVTKPK
ANANKQONTPAKARKQDAVAAANMPGSENNIDILRLAQVYGEAREPEECOVATLAAV
ILNRLNSPLFPNSVAGVIFEPLAFVADGQIYQMDNETAREAVDAINGMDPSEBAL
YFNPPTATSPWIMGRPQIKRIGHICE"

gene

complement(4162. 4841)

terminator

/gene="ypdC"

CDS

complement(4162. 4841)

terminator

/gene="ypdC"

complement(4185. 4841)

/gene="ypdC"

/function="unknown"

/note="similar to hypothetical proteins"

/codon_start=1

/transl_table=1

/protein_id="CAB14226.1"

/db_xref="GI:2634729"

/db_xref="SWISS-PROT:P50738"

/translation="MFATISAGIAPGIALLSFYLYLKQDYDNEPVHMYLRSFFLYLV
FPMFQYVLEKENVGGSFVSWSGFLESKLWFIIMISVYPAHFDEYDGVY
GASVSGFATLENILLYLIGHGVHAFRALLPVSCALIGVMGYLGAARSSADAKR
VWMLTSLVPSLHGSYDFILTALSNMIYMLPFVFMFGLRKAKKARSVMNQV"

gene

complement(4929. 5935)

terminator

/gene="ypdA"

complement(4929. 4953)

/gene="ypdA"

complement(4961. 5935)

/gene="ypdA"

/function="unknown"

/note="similar to thioredoxin reductase"

/codon_start=1

/transl_table=1

/protein_id="CAB14227.1"

/db_xref="GI:2634730"

/db_xref="SWISS-PROT:P50736"

/translation="MIOGKAIIIGGPGCLSAIHLKQIGIDALVIEKGVNYSIYN
PHQTFESSSEKLEIGDAFITEENRPVPIQALSYREVYKRNIRVNAFEVNRVTK
TONNTFVIEITSEKETYTPYCIATGYDHPNMGVGEDLPVFNHFKGHYFDKDV
VWIGKSSVDALDELKSGARVTVLYRGNEVSPSIPVILPEFEALVANGTIRMEFG
ACVEKITEVNEVRSGEKELITIKNDVFPAMGYHDDHOLEKIGVEIDKETGRPEFN
EEMETNVGCVFIAGYIAGNANNEFIENGCFHGHIAELAKKENH"

terminator

complement(5995. 6016)

gene

/gene="ypcA"

complement(5995. 7324)

/gene="ypcA"

complement(6044. 7324)

/gene="ypcA"

/function="unknown"

/note="similar to glutamate dehydrogenase"

/codon_start=1

/transl_table=1

/protein_id="CAB14228.1"

/db_xref="GI:2634731"

/db_xref="SWISS-PROT:P50735"

/translation="MAADNTGHTEDKLDVLKSTQTVYHKLKLEKLGIPEDVYELLE
PKRLTVKIPVRDDGSVKIFGTGRAHDSVGTGIGIRFHPNVTKEVEKAVKALSIW

gene
complement(2134. .3552)
/gene="ag_128"
/complement(2134. .3552)
/gene="ag_128"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAC06502.1"
/db_xref="GI:2982882"
/translation="MORIIIVNPEPYLSVIRKRVKLSIPITIVNLLYVENNISMLV
SSISPAVAATFSLTLMFTISLMAISYSTNIIIAQVGAKDPSILINGLPSF
LISLPEYFGKDFVFLMKVIGASETVRSIAKEYLPIEMFPIGELNTEFYGANCA
GDTKPMKVAIITMNLTHGTATLNGKRGALGEGAGWGIAISEILAFIYFELL
IFKKPPLHLALPEPKLIFKMYRLCTPALREAIITLSPNVYFGFLAKGDKVLAHO
IGRLTESIFMIGFGVMIASTTLAAGONGARYRMVHNVNTSAHFTLVMSLTCLIL
ILFPHYLVYPSRDEVEIEMASYLQIVGISQPAAYASISGALKGKGNKHIPLVN
ISSFWLRIIPSYFLKVIHSPVLVPMGWTFTETAVRALFYTVFKKVVYVKLQTPQKE
EEPEGDEAGDPLDEVRSKYV"
/gene="ag_131"
/note="ribh"
/gene="ag_132"
/note="ag_132"
/db_xref="GI:2982882"
/gene="ribh"
/codon_start=1
/transl_table=11
/product="riboflavin synthase beta subunit"
/protein_id="AAC06489.1"
/db_xref="GI:2982869"
/translation="MOIYEGKLTAEGLRFGIVASRPHNALVDRLEGAIDCIVRHGGR
EEDILVHVPGSWEIIPVAAGELARKEDIDAVIAIGLVATPHFDYIASSEVSKCLAN
LSLELRKPTTGVITADPLQOALIERAGTKHGNKGMALSAIEMANLTKSLR"
/gene="ag_133"
/note="nusb"
/gene="ag_133"
/note="ag_133"
/db_xref="GI:2982871"
/translation="MRYRGARDTAEFLVLRWMDLGENGELFEKEVEENKINKNDAY
EYAKLVDTAVHIEIDSLIEKHLKSGSIDLQAGVARNALIGVAELLFLKSKPEGR
VFIDVLDVYKXADKAGKRVNGVLSAITYKATTSKEKPSLASE"
/gene="ag_134"
/note="ag_134"
/db_xref="GI:2982880"
/translation="MROYSYTGSSKASEEVEFAVRLCKELAKRNLVYVCGGTGVM
AYCKAKKEGGGLTIGIMPSYDSGHEANPYIDIVYNNMNNRPYVASSEMVLAAGN
IGTSLIAYALLGKIYVGYKTHKVEGVQVESDEMLKADVEFGRL"
/gene="nueh"
/note="ag_135"
/db_xref="GI:2982870"
/translation="NKVFITGATGFGVGRHIVELLNRGYEYHAGVRNLSKLERLFGNO
VKGIYVNDKDSIREALGKVPDVIVHILGILYEKKGTTFEVRHGHKNLVEVS
KGFNKKRFLFMSALGTDEAPSRHQTRMARREVINGLANTTIERPSILIGPEKLF
FDMYIKTVIPIVVALPDGNGVQFOPVDVADVACAALAKNETRKLKYEILCGTAVV
FKELADIYSYDRKVLMIIPMKIKMYTGLVETIIEPPSPSSQIKMMKPNVCGV
LKDEAISNGIKSICKRPIPYEESLKMSENLDIILKKK"

gene
complement(7162. .8232)
/gene="cpx"
/note="ag_136"
/db_xref="GI:2982865"
/gene="cpx"
/codon_start=1
/transl_table=11
/product="cytochrome c peroxidase"
/protein_id="AAC06485.1"
/db_xref="GI:2982865"
/translation="WRKGLIMAVFAGTVAKKIDDKELIKMARQIKPLPKAENP
QNPVPEKVKIGKMLYDPRLSGLJSCNCHNARVGNLPTSIHRVAIGRMA
PTVVAALHIOFMDGARAKDVEQALIVNPIEMANTEENAVNLTKISIPYVELFK
AEPNKDVKKENIGKALGARERLMPSRDEFKGNTRKALTDQERKGLFTIEVGC
VACHNGPVGNGMTRKPMGITIEYKVPYLYVQKPAIKVDFGREGVTKREDDVEK
VPSLNIHHTPYRHDSVMSLEDVAVRIMATQLOKELTDQVADYAFALALGKIP
KHALEVPPLPSTIDTPKPEL"
/gene="ribd1"
/note="ag_138"
complement(7162. .8232)
/gene="ribd1"
/codon_start=1
/transl_table=11
/product="riboflavin specific deaminase"
/protein_id="AAC06487.1"
/db_xref="GI:2982867"
/translation="MIREIDNVMKLAISLAKKRRGYTHNPYGVAVYKRGKIVGLG
YHERAGKHAEMVMAIQAGEKAKATLVLTPECTHFGRTPTCDALIRSGIKRVA
LIDNPMLSGGVEKLRNAGIEVVGVEEBAERLEDEFTYITQERYITLKAOTL
DGKIALTLGSSKWIISKSRKVAHLIREATAVAVGNTVTKDPHLVRFVPEKOP
VRILIDPELEPLSAKVLNTEAPAVITVTKENKRAKELKRGVOVLIKGFNLKIL
KKKLEIMHLMVREGGPRITLSFLKEGFPDRIVYFIARKINGEGLSIDDLRISEES
LKVRKKVENLGEDVLVIFKRY"
complement(8229. .9173)
/gene="ribf"

gene
complement(7162. .8232)
/gene="ribd1"
/codon_start=1
/transl_table=11
/product="riboflavin specific deaminase"
/protein_id="AAC06487.1"
/db_xref="GI:2982867"
/translation="MIREIDNVMKLAISLAKKRRGYTHNPYGVAVYKRGKIVGLG
YHERAGKHAEMVMAIQAGEKAKATLVLTPECTHFGRTPTCDALIRSGIKRVA
LIDNPMLSGGVEKLRNAGIEVVGVEEBAERLEDEFTYITQERYITLKAOTL
DGKIALTLGSSKWIISKSRKVAHLIREATAVAVGNTVTKDPHLVRFVPEKOP
VRILIDPELEPLSAKVLNTEAPAVITVTKENKRAKELKRGVOVLIKGFNLKIL
KKKLEIMHLMVREGGPRITLSFLKEGFPDRIVYFIARKINGEGLSIDDLRISEES
LKVRKKVENLGEDVLVIFKRY"
complement(8229. .9173)
/gene="ribf"

Query Match 11.3%; Score 116.4; DB 1; Length 14364;
Best Local Similarity 50.8%; Pred. No. 3.7e-12;
Matches 386; Conservative 0; Mismatches 361; Indels 13; Gaps 4;
QY 1 ANGATTTATCGATTCACCTTGCAATATGTCACAGTCAACAGCGTGTATCCACC 60
DB 8205 ATGAAGTTGGACCTCTCTGCTGAAGAGAAAAGGCTACATCATCTCAAGT 8146
QY 61 GTTGGCGCTGTTGATTAATGAGGTAGCATTTGTTGTTGTCACACTTGAAGAAA 120
DB 8145 GTGGGAGCGGTGTGCTAAGAGAAAACGTGGGATTAGGTTATCCAGAAAAGCA 8086
QY 121 GGTGACACAGCAGCGAGGTCAACGACTGATATGCGACAAACAAATGCGAAGTGG 180
DB 8085 GGAAGCCTCAGCGTGAAGTATGCTCGGTCAAGCGCGGGGAAAACAAAGGGGCA 8026
QY 181 AGCATTTATATACGTTAGAGCCATGATGCAATTTGTTGTTCAACACCAACCTGTGTAAC 240
DB 8025 ACCCTTACGTTACCTTACCGCTGACCGGTGATCTCACTTCGGAGAACTCTCCACAGGAC 7966
QY 241 AAAATTATTGATTGTAAGATAGCAAAAGTAGTATACGCAACAAAAGAC-----AATTGC 294
DB 7965 GCGATATATAGAGCGCGGATTAAGAGGTTGATGCGCAACCTTGACCACCAATCCTTTG 7906
QY 295 TTAGACACACATGCTGATGAGACGTTAGCGGCTCAGCGATATGAGTGTGAATGCGTTG-- 352
DB 7905 ATGCTCTGGAAGAGGTGTGAAAACTGAAAAATGCGGGAATVGAAGTGATGTCGGTGA 7846
QY 353 -ATGATGAACGGGCAACCAATTATACCAAGACTTTTAAAGCAAAAGCAAGCAACTG 411
DB 7845 TCGGAAGAGAGAGCCAGAGAACTCAACGAAGATTTTTTTACTATACATATACGAGGA 7786
QY 412 CCACAAATTACAGTGAAGATATCTCAAGTTAGATGTTAAACAAGCAAGATATATGGA 471
DB 7785 CCTTACATTAACCTTAATAATGAGCCAGACCTGAGACGGAACCACTCGCACCTCAGCGGA 7726

QY	472	CANSTCAATGGATTACTACAAACAAAGAGTTTAAACGAAGTGTCTATTAAGTAAAGCAATTCGA	531
Db	7725	AGCAATAGTGGATTAACCTCAAGAGGAAGCCAGAAAGTAGTCTACATTTTAAAGGAGAG	7666
QY	532	CACGACGAGTGTTAACGGAAGAGTACAGTGAATTAAGATCCACAAATATATCTACA	591
Db	7665	GCAACTGGGTTCTGTAGTGTATAACACAGTCATTAAGAGACCCACACTTAAACGCTC	7606
QY	592	CG---TATTCAAGATGAAAAAACCTTATAAAGTAAATATTGTCTAGTCTGGGAAATTT	648
Db	7605	AGGTTGTCTCCAACGGAACCAACACCTGAAGATTAATCTGGACCCCTGAATCTGAAAGTT	7546
QY	649	CATTATATCAACCAATTTATCAACATATATCAACACCAATTTGGATATATACAGAAAT	708
Db	7545	CCACGAGTGTCTAAAGTTCTGAATACGGAAGAGTCCACAAANTTTCTATTAAGCAAAA	7487
QY	709	CCAAATTTACAGCAATCAACACATATTTGAAATATTTT	748
Db	7486	GGAGAAATGAAAAAGCTGAAAAAACTCAAGAAAAAGAGATT	7447
RESULT	10		
BARIGENS			
LOCUS	4286 bp	DNA	07-AUG-1997
DEFINITION	B.amyolliquetaciens ribb, ribc, riba, ribh & ribt genes.		
ACCESSION	X95955		
VERSION	X95955.1		
KEYWORDS	GI:1592687		
SOURCE	deaminase; GTP cyclohydrolase II; imazine synthase; reductase; ribA gene; ribB gene; ribG gene; ribH gene; riboflavin biosynthesis operon; riboflavin synthase; ribT gene.		
ORGANISM	Bacillus amyolliquetaciens		
REFERENCE	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillaceae; Bacillus.		
AUTHORS	1 (bases 1 to 4286)		
TITLE	Gusarov,I.I., Kreneva,R.A., Podchernyaev,D.A., Iomantas,Y.V., Abdlakina,E.G., Stoinova,N.V., Perunov,D.A. and Kozlov,Y.I.		
JOURNAL	Riboflavin biosynthesis genes of Bacillus amyolliquetaciens: primary structure, arrangement and regulation		
REFERENCE	Mol. Biol. 31, 446-453 (1997)		
AUTHORS	2 (bases 1 to 4286)		
TITLE	Gusarov,I.		
JOURNAL	Direct Submission		
COMMENT	Submitted (27-FEB-1996) I. Gusarov, Gnitgenetika, 1st Dorozhny pr. 1A, 113545, Moscow, RUSSIA		
FEATURES	On Oct 4, 1996 this sequence version replaced gi:1212772.		
source	Overlaping sequence 227458.		
location/Qualifiers			
1..4286			
/organism="Bacillus amyolliquetaciens"			
/strain="A 50"			
/db_xref="taxon:1390"			
321..348			
/note="riboflavin operon"			
653..1768			
/gene="ribc"			
653..1768			
/gene="ribg"			
/codon_start=1			
/transl_table=1			
/product="geaminase"			
/protein_id="CA65189.1"			
/db_xref="GI:1592688"			
/db_xref="SWISS-PROT:P70814"			
/translation="MEEYVNTATLELARRGEGOTNPDLVAVVYKKRQIVGMAHLD			
YGEAHVHAHINAGSLAKGADIVTLTEPCSHYKCPKPAELIMKSGIKRFLAVEDP			
NPLVAGKITYLEKAGILEVTKGILRQQAEBELNKFELPMRTGELFYVTLKAAASLDGT			
ATFEGDSKWIITSEARLDAQOYRSHORILVAGTVKADNPSTLCRLPDVAKQIPRYI			
LDLTALPETERANVLTGDAAPFTWITFAGSDVRRKDRFALGAIKVFTEIETRIHPEVL			
SILKENGIMSVYEGGASVHGSEFVAKCFDELHLYFPADIIIGTILAPSLISGEGFSM			
KDVPHLOFTOTIGTQIPDKILAIKPKDKDGDVYRN"			
1749..2396			
/gene="ribb"			

```

CDS      1749..2396
        /gene="ribn"
        /codon_start=1
        /transl_table=11
        /product="riboflavin synthase (a-subunit)"
        /protein_id="CA65190.1"
        /db_xref="GI:1212774"
        /gene="ribn"
        /note="GTP cyclohydrolase II"
        /codon_start=1
        /transl_table=11
        /product="3,4-dihydroxy-2-butanone 4-phosphate synthase"
        /protein_id="CA65191.1"
        /db_xref="GI:1212775"
        /db_xref="SWISS-PROT:P51695"
        /translation="MFPIEEALEAKKGEYIIIVDDEDERNEGEFVALAEHATPEVW
NEMTHRGRLICPLSDIEDIGRLDHPMVNHTSHEFAFVSIDHKLTGTGISAOER
SEITQALDDSESVGPDRPGHTEPPLIAKKGGYVKRGAHPAANDLAKCSOGADVI
CEINEDGTAARVPETSEIASKSHOLKAITTKDLIEKYNTITLVREVDITLPDEFST
FRVYGTVNEVDGKEHLAFVMGDVEFNSEPVAVRVHSCLTGDVFASHRCDCGPOLAAH
LRJLAEGREGVLTLROEGRGIGLINLKAYRILOEGYDVVEANEALGFIPDLNVGIGI
GAOILRDLOGVQHMKLTNNPRKIAGLEGYGLSISDRVPLQMEASEHNKOYLQTMMKRL
GHLLHF"
gene     3639..4103
        /gene="ribn"
        CDS    3639..4103
        /gene="ribn"
        /codon_start=1
        /transl_table=11
        /product="lumazine synthase (b-subunit)"
        /protein_id="CA65192.1"
        /db_xref="GI:1212776"
        /db_xref="SWISS-PROT:Q44681"
        /translation="MNIIGNLVGTGLIGIVGRENFETYSKLLSGAEDTLIRHGYE
SNIDVAMVGGAPEIIPPAKKMAETKLYDAVITLGTVIRGATTHDYVCNEAKGIAD
AGRTGVPIVPEGIYTITIEDAIRACTKNKCADCANVAIEVANLNRSFE"
gene     4211..4286
        /gene="ribt"
        CDS    4211..4286
        /gene="ribt"
        /codon_start=1
        /transl_table=11
        /product="reductase"
        /protein_id="CA65193.1"
        /db_xref="GI:1212777"
        /db_xref="SPRMBL:Q44682"
        /translation="MLIRRYKKSFEKIMGLSFMPNE"
BASE COUNT   1306 a   926 c   1102 g   952 t
ORIGIN
Query Match          10.9%; Score 112.2; DB 2; Length 4286;
Best Local Similarity 52.0%; Pred. No. 2,5e-11;
Matches 305; Conservative 0; Mismatches 273; Indels 9; Gaps 2;
QY       1 ATGATTATCGCATTCACCTTGCAAAATATGTACAAGTCAAACAGGTGTTAATCCACC 60
         ||| || | || | || | || | || | || | || | || | || | || | || |
Db        668 ATGATTAGGCACATAGACTCTGCCAGCGCGGTGAAGAAGACACAGCCGAAATCCTCT 727
QY       61 GTTGCCGCTGTTGTAAGTAAATGAAGCTAGGATTTGTTGATATGGTGCACACTTGAGAAA 120
         || |||| | || | || | || | || | || | || | || | || | || | || |
Db        728 GTCCGCGCCGTCGTGTCATAAAAAAGACAGATCGTCGGAATAGGGGCCCATCTTGAAATAC 787
QY      121 GGTGACAGCATCGCGAGAGTTCAACCACTTGATATGGCACAAACAAATGCTGAAGTGC 180
         || || | |||| | |||| | || | || | || | || | || | || | || | ||

```

Db 788 GGAGAGGCCATGCTGATGTCACGCGATCAATATGGCAGGACCGCTAGCAAAAGCGCT 847
Qy 181 ACATTAATATTAAGCTAGAGCCATAGTATGTTGTTCAACACACCGCTGTATAC 240
Db 848 GATTGTATGTACGCTTGAGCCCTGACGCAATTAAGCAAAACACCGCATGTGCCGA 907
Qy 241 AAAATTAATGATTAAGATAGCAAAAGTATACGCAAAAG-----ACATTCG 294
Db 908 CTGATCATGAATCAGGAGATAAAAGGTTTATTCCTGTGGAAGATCCCAATCCGCTT 967
Qy 295 TTGACACACATGTGTATGACGCTTACGGGCTCAGCGTATGAGTTGAAT---GCGTT 351
Db 968 GTAGCCGGAAGAGCATCACATGCTGGAAGAGCGGCATCGAAGTGAAGAAACGGGCTT 1027
Qy 352 GATGATGAAGGGCATACATTAATTAACAGACTTTTAAAGCAAAAGCAACCACTG 411
Db 1028 CTCAGGCAACAGGAGAGAGCTGAATAAATGTTTCTTCATTCAGCGGACCGGACTT 1087
Qy 412 CCACAAATTAACAGTGAAGATATCTGAAGTTTATGATGTTAAACAGGAAATGATATGA 471
Db 1088 CCGTATGTCACCTTGAAGAGCGCGCAAGCCTTACGCGGAAACAGGACGATGAACCGGT 1147
Qy 472 CAAATCAATGATTAATTAACAAAGAGTTAAACAGATGTTATTAAGTAAAGATGA 531
Db 1148 GACAGCAATGATTAATTAATCAAGAGCGCAAGGCTTATGATCCCGCATTAAGAGAAATCG 1207
Qy 532 CAGAGCGCATGTTACTGGAAGACGTACAGTTGATTAATGATATCC 578
Db 1208 CATCAAGATGTTCTGTGAGAGCAGGACAGTCAAAACCGATATATCC 1254

RESULT 11
LOCUS U32775 18073 bp DNA BCT 29-MAY-1998
DEFINITION Haemophilus influenzae Rd section 90 of 163 of the complete genome.
ACCESSION U32775 L42023
VERSION U32775.1 GI:1573951
KEYWORDS
SOURCE Haemophilus influenzae Rd.
ORGANISM Haemophilus influenzae Rd.
Bacteria: Proteobacteria; gamma subdivision; Pasteurellaceae; Haemophilus.

REFERENCE 1 (bases 1 to 18073)
AUTHORS Fleischmann,R.D., Adams,M.D., White,O., Clayton,R.A., Kirkness,E.F., Kerlavage,A.R., Bult,C.J., Tomb,J., Dougherty,B.A., Merrick,J.M., McKenney,K., Sutton,G.G., Fitzhugh,W., Fields,C.A., Gocayne,J.D., Scott,J.D., Shiley,R., Liu,L.I., Glodek,A., Kelley,J.M., Weidman,J.F., Phillips,C.A., Spriggs,T., Hedblom,E., Cotton,M.D., Uterback,T., Hanna,M.C., Nguyen,D.T., Sauder,D.M., Brandon,R.C., Fine,L.D., Fritchman,J.L., Fuhrmann,J.L., Geoghagen,N.S., Gnehm,C.L., McDonald,L.A., Small,K.V., Fraser,C.M., Smith,H.O. and Venter,J.C.
TITLE Whole-genome random sequencing and assembly of Haemophilus influenzae Rd
JOURNAL Science 269 (5223), 496-512 (1995)
MEDLINE 95350630
REFERENCE 2 (bases 1 to 18073)
AUTHORS Tatusov,R.L., Mushegian,A.R., Bork,P., Brown,N.P., Hayes,W.S., Batousov,K., Rudd,K.E. and Koonin,E.V.
TITLE Metabolism and evolution of Haemophilus influenzae deduced from a whole-genome comparison with Escherichia coli
JOURNAL Curr. Biol. 6 (3), 279-291 (1996)
MEDLINE 96396784
REFERENCE 3 (bases 1 to 18073)
AUTHORS White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.
TITLE Direct Submission
JOURNAL Submitted (25-JUL-1995) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA
REFERENCE 4 (bases 1 to 18073)
AUTHORS White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1997) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA

REMARK The H. influenzae sequence has been updated by R. Fleischmann. New database matches have been assigned, product names have been improved, and a number of frame shifts have been corrected. We gratefully acknowledge the work of Tatusov et. al. We have incorporated their annotation into the /notes fields of the corresponding H. influenzae genes

REFERENCE 5 (bases 1 to 18073)
AUTHORS White,O., Clayton,R.A., Kerlavage,A.R., Fleischmann,R.D., Peterson,J., Hickey,E., Dodson,R. and Gwinn,M.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-1998) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA
REMARK The whole genome was shifted by 588 nucleotides for a new start on Sep 30, 1996 this sequence version replaced gi:1221677.
COMMENT
FEATURES
source
1..18073
/organism="Haemophilus influenzae Rd"
/db_xref="taxon:71421"
/complement(146..1327)
/gene="HI0929"
complement(146..1327)
/gene="HI0929"
/note="similar to GB:M77129 SP:P24196 PID:146678
PID:882568 GB:U00096 percent identity: 41.10; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC22589.1"
/db_xref="GI:1573952"
/translation="MKRVGFPPTPDVYQOQLNVGFDDYVNLPSDGSNHYSDNVAYE
FLAEIDRIEDTNNELSMCLDFADEIKKGDENVHFELOKLETSWRNDEPLVY
RDRGVDGNNKMFENYADTPSLLEPAVYQWOMEQEGKHPDQKPIHLEIKHP
QFLKQSGKTDHLSAMODAGREDNGVNDYLDVYVNGMNIHQLAVEDIGNSTK
FVLDNDQPLEMLFKYPLEWSHAEFARHTTAETRFLEPAKMLLSNKAALAKLMAR
YNNHDLPAVTFPELPRDLMSVMWKKPLGREGANVYEEKNNQVNERPAKGSERST
YGNSHVYVQKKEFLPSFDGMYPIIGSWVDVACGMGLREDFTAVTGNDSHPIPHYFV
P"
complement(1328..1948)
/gene="HI0930"
complement(1328..1948)
/gene="HI0930"
/note="hypothetical protein; identified by Genemark;
putative"
/codon_start=1
/transl_table=11
/product="H. influenzae predicted coding region HI0930"
/protein_id="AAC22600.1"
/db_xref="GI:1573966"
/translation="MARRKNHSIKIENNEKENOILVSLIALLGCGSEGOYORDVQ
SIDDLAAMKKELEADKNPESITOKTATTOOQGLNRDNGMLESVKNPAENNAQ
AKQSEKRAESTTKAESTDPSLGAATAGVITGMARALISFEGPEYHGNRAVTPPG
QVVOPTNRSVCKPMLVGNAGSNKSPVSRGFSPPKTRHSSG"
complement(1921..2406)
/gene="HI0931"
complement(1921..2406)
/gene="HI0931"
/note="hypothetical protein; identified by Genemark;
putative"
/codon_start=1
/transl_table=11
/product="H. influenzae predicted coding region HI0931"
/protein_id="AAC22601.1"
/db_xref="GI:1573967"
/translation="MKSRTLYKSAVFLYLFTKGRMLHLLEDFLGTCKPQVSTHS
TVFEHFAVWFEDDEGTGYFALDMRONKOPIDVMDHVNVDSTSNHBEARKLEICWDE
SGYVALLLINGPFAVDFEARLVGNSKRYPOPDLMSWMTREITNKQAEQMLGKTI
R"
complement(2424..3734)
/gene="HI0932"
complement(2424..3734)
/gene="HI0932"
/note="similar to GB:M12843 SP:P08324 GB:X82400 PID:147479

gdmlyvtgdlgdsqagldvildpekrhl.pfadiileorhlystprlvaqoalvhlssa
idsldgladiolhllrrsnvgsidvslpkslelgojvdsstaojyaltsgeeyel
cfttpebnrgslennlahcgtvvtcigolrpragtelhnoqkklmqlagldhfevka
complement(7014. 7484)
/gene="VC2267"
complement(7014. 7484)
/note="VC2267"
/note="similar to GB:M26839 SP:P04381 GB:X0681 GB:X64395
PID:146976; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="N utilization substance protein B"
/protein_id="AAF95411.1"
/db_xref="GI:9656829"
/translation="MGASYKPARARNARQFALQAIYSMTKENVATIEOFTSGXY
DEEHRAPALAPETDVSFRDLGAVLHNELDSLRFVFRPPODDMELAL
LRLAYEMTRRDVPRKVINAEIELAKVFAEDSHKFEVNGVLDRAAPHRKKA"
complement(7484. 8005)
/gene="VC2268"
complement(7484. 8005)
/note="VC2268"
/note="similar to GB:L11391 PID:496182 SP:P51963;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="riboflavin synthase, beta subunit"
/protein_id="AAF95412.1"
/db_xref="GI:9656830"
/translation="MLESQSHLMNRVKGKLMKRVIEGGFPAPNARAIIVISFNSFIN
ESLSGADIDTLKRHGOISPDNITVRCBPVAVLPVAGVATGDDIVISGCVIRG
GPHFDYCEENKNGIAQVSLFSPVAFGVLTVDIQAIRAGTKANGKGAEMALS
ALEMNVLSSEIDS"
complement(8219. 9328)
/gene="VC2269"
complement(8219. 9328)
/note="VC2269"
/note="similar to GB:M90094 SP:002008 PID:150708;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="3,4-dihydroxy-2-butanone 4-phosphate
synthase/GTP cyclohydrolase II"
/protein_id="AAF95413.1"
/db_xref="GI:9656831"
/translation="MPSIPQDIIEDIRQGRKVIIMDDEDRNEEDLIIMAHITPAA
INFMATHGRLICLTIFKRCRRLGLNPMVQDNNAQYTNFTVSTIEAEGVTTGISAA

Query Match 8.9%; Score 91.4; DB 1; Length 11695;
Best local similarity 49.7%; Pred. No. 1.4e-07;
Matches 292; Conservative 0; Mismatches 286; Indels 9; Gaps 2;

QY 1 ATGGATTATGCGATTCACCTGCAAAATATGTCACAAAGTCACAGGTGTTATCCACC 60
DB 11097 ATGCTCGCGCGCAATTCGCTGCGCGCGGCGGCTTTACCACTTCCTCTAAT 11038
QY 61 GTTGGCGGCTGTTAGTATGAAGTAGGATGTTGTTAGTGCACACTGAGAAAA 120
DB 11037 GTGCGGTGGTGATCACTGCGGTGAGACAGATTGAGGGAAGTTCCATTTCCGCG 10978
QY 121 GGTGACACAGCATCGAGGTCAAGCACTGATGCGACACAAATGCTGAAGTGGC 180
DB 10977 GCGGAACCCATCTCTAAGTCATGCTATGCGCAAGCCGCTTACTCGCGGTGCG 10918
QY 181 AGCATTTATATACGTTAGACCATGATGATTTGGTTCAACACACCCTGTGTTAC 240
DB 10917 ACCGCTTATGCTCTTAGAGCTGCTCATTAAGTGCGACACCGCTGTGCTGAA 10858
QY 241 AAAATATTGATTTGAAGTAGCAAAAGTAGTATAGCAACAAAGACAT-----TGC 294
DB 10857 GGGCTTTATTAAGCCGGGGTTGGAAAGTAGTATTTGGCATGCAAGACCTTAACCA 10798
QY 295 TTGACACACATGATGATGAGAGCTTACGGGCTCAGCGGTATGAGTGAATGCGTTGAT 354

Db 10797 GTGGCGGGGAAGCGGTACCAATGCTGCGTATGCCGGATTGAGGTGAAAGTGGGTTG 10738
QY 355 GATGAACGGGCACTCACATTA---TACCAAGACTTTTAAAGCAAAAGCAAGCAACTG 411
DB 10737 CTGAAGCGGATGCTGAGTACTCAACCGCGTGTTCCTTAAACGATGAAACCGGCATG 10678
QY 412 CCACAAATTACAGTGAAGTATCTGCAAGTTAGATGCTAAACAAGCAATGATATGA 471
DB 10677 CCTTTTGTCACACTTAATATGCGCGAGTCTCATGTGACAAACCGCCTGCTAACGCT 10618
QY 472 CAAGTCAATGATTAAGTACGTAACAAAGAGGTAAACAGATGCTATAGTTAAGACATGA 531
DB 10617 AAAACCCAGTGGATTACCTCTCTCCACAGCTAAAGATGTTACAGCGGTTTCGCCGCAA 10558
QY 532 CACGACGACGTGTTAAGTGAAGAGCTAGTGAATTAAGTATGAC 578
DB 10557 GCGAGTGCATTCCTCCACACGATCAACGAGTGTGCGGATATGC 10511

RESULT 13
AE001819/c
LOCUS 12811 bp DNA BCT 02-JUN-1999
DEFINITION Thermotoga maritima section 131 of 136 of the complete genome.
ACCESSION AE001819 AE000512
VERSION AE001819.1 GI:4982396
KEYWORDS
SOURCE
ORGANISM Thermotoga maritima.
Bacteria: Thermotogales: Thermotoga.
REFERENCE
1 (bases 1 to 12811)
Nelson,K.E., Clayton,R.A., Gill,S.R., Gwinn,M.L., Dodson,R.J.,
Hart,D.H., Hickey,E.K., Peterson,J.D., Nelson,W.C., Ketchum,K.A.,
McDonald,L., Uitterback,T.R., Malek,J.A., Linher,K.D., Garrett,M.M.,
Stewart,A.M., Cotton,M.D., Pratt,M.S., Phillips,C.A.,
Richardson,D., Heidelberg,J., Sutton,G.G., Fleischmann,R.D.,
Eisen,J.A., Fraser,C.M. et.al.
Evidence for lateral gene transfer between Archaea and bacteria
from genome sequence of Thermotoga maritima
Nature 399 (6734), 323-329 (1999)

TITLE
JOURNAL 99287316
MEDLINE
REFERENCE
AUTHORS
2 (bases 1 to 12811)
Nelson,K.E., Clayton,R.A., Gill,S.R., Gwinn,M.L., Dodson,R.J.,
Hart,D.H., Hickey,E.K., Peterson,J.D., Nelson,W.C., Ketchum,K.A.,
McDonald,L., Uitterback,T.R., Malek,J.A., Linher,K.D., Garrett,M.M.,
Stewart,A.M., Cotton,M.D., Pratt,M.S., Phillips,C.A.,
Richardson,D., Heidelberg,J., Sutton,G.G., Fleischmann,R.D.,
White,O., Salzberg,S.L., Smith,H.O., Venter,J.C. and Fraser,C.M.
Direct Submission
Submitted (01-JUN-1999) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
Location/Qualifiers
1..12811
/organism="Thermotoga maritima"
/db_xref="taxon:2336"
172..591
/gene="TM1815"
172..591
/note="TM1815"
/note="similar to GB:L77117 SP:Q57699 PID:1499030 percent
identity: 60.56; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="Ferredoxin"
/protein_id="AAD36878.1"
/db_xref="GI:4982397"
/translation="MAEAKNAPLLGKDALGREVDSKVPWVGDRKEIEMYPPIYD
KCVGCGICFVTCGRRVDFDKKEGVIVARVNCMAQCTCMNLCPTGATISFPDASTI
KKLVAOKRIYKAEFIILPLAEDHLSPKETETPP"

gene
CDS

```
/gene="TM1816"  
/note="similar to GB:Pyro.h percent identity: 65.04;  
identified by sequence similarity; putative"  
/codon_start=1  
/transl_table=11  
/product="conserved hypothetical protein"  
/protein_id="AAD36879.1"  
/db_xref="GI:4982398"  
/translation="MIIAPVSENGKSPISSEHGRRPYAFVYKNNAIADISVEE  
NPADQDHGAVPNFKERKAGALVIRIGRRAIAPFAMGVKVIKASGIVEEVNQ  
VLSQGLKDSDEYEHDDHHHHH"  
/gene="TM1817"  
complement(1011..3608)  
/gene="TM1817"  
complement(1011..3608)  
/note="similar to GB:M16318 SP:P11931 PID:143797 percent  
identity: 72.43; identified by sequence similarity;  
putative"  
/codon_start=1  
/transl_table=11  
/product="valyl-tRNA synthetase"  
/protein_id="AAD36880.1"  
/db_xref="GI:4982399"  
/translation="MAELSTRYNPAIEITKMYRWEKGYFTPKGCEKFSIYPPN  
ITGRIMHGALNITLQDIYVRKRMKGYDLAVSPEDHAGIATONAVEKFLQTOGK  
REIEGREKFLIETWEMANKYREIREIOIKALGASVDMTRERETLDEGLSRAVRKVE  
LYRKGLIRGKTIYVWCPRCKTVLSDEVEHKEHKSILYKRPKDSDEIVYATTA  
PEMLGDTAVAHPEDEKTKNFVGTLLPLVGREIPVADYVDPKGTGAVKATPA  
HDPNDYVLAQRNLPMIEIFDNRINNGKYGKGLDYERAEKIVKDLGOGVLKI  
EDYTHSGHCRDVTIEPKLSDQWVSTKPLAKRAIYAVENGELRFPPEMTKYLN  
MWYELRDMCISQLMWGRIPWYQCDCGHLNASEDYKCEKCGSTNLKODEVDLT  
MSSALPDCSTLGMPEETEDLKRYPTDLYVGPDIIFFWARMTMGVEFEMNDKPS  
HYITIQIQRDKYGRKMSKSLGNDIPLEVIDEYGDMPRFTLALIAAGROIKDPRY  
DFAIKRFKRNATREYVMDNEDKEYVLEMLKTVDMWILFRKTYEETYNALEN  
DFNIAARTIYFEMDFCDWYIEASKPRLKTEERNLVQTVLVKVLDAASLYRLHPMP  
LTELMOKLPVAGESITIAKWPETEIRELIDETAKEKEFTRLMMVGVNRVNAENMLPQ  
SORVYVYIKGEVTEEBELLTKLTGNIIEVSFVNEKPKPTATAYEEIEIYVDIGLT  
IDFEKERLRKQIMEIKQEIIDLREKLANDKDFVAKPEVEVEETKEKLNRRRLAR  
LESTIRDE"  
complement(3640..4140)  
/gene="TM1818"  
complement(3640..4140)  
/note="similar to GB:AE000783 percent identity: 61.07;  
identified by sequence similarity; putative"  
/codon_start=1  
/transl_table=11  
/product="chromate transport protein, putative"  
/protein_id="AAD36881.1"  
/db_xref="GI:4982400"  
/translation="MLKLAFFIKIGLSPFGGNAIVGILKNELVYTGSEFSPESP  
VSIADMPGPVAINLATITGKFGILGAVLNTLFLDAPLIVITTAFLKRYKLR  
VRLKALGATTTLLIVTLSSLSSVNPVILLSSAAVFCSEFKVHPLFIIFGCVI  
GAILGF"  
complement(4144..4716)  
/gene="TM1819"  
complement(4144..4716)  
/note="similar to PID:687689 percent identity: 55.06;  
identified by sequence similarity; putative"  
/codon_start=1  
/transl_table=11  
/product="chromate transport protein, putative"  
/protein_id="AAD36882.1"  
/db_xref="GI:4982401"  
/translation="MEICVEGSKSRIEVEQSNLIGLFLFLRISALTIGGYAMIP  
VMKLEERSGLTEKEFRIMSTAOVPGPIAFNTAVVGRRLAGIYGAISGLAVVL  
PPFAIVAVAEVIRLIGSIYSVRSFRLGAVASIIGLVGSVLYRLVRNQRMVLYRAIMI  
GVAVFVLLNSVIVPVLVLLVLLKEV"  
4772..6277  
/gene="TM1820"  
4772..6277  
/gene="TM1820"
```

```
/note="similar to GB:AE000657 percent identity: 76.48;  
identified by sequence similarity; putative"  
/codon_start=1  
/transl_table=11  
/product="GMP synthase"  
/protein_id="AAD36883.1"  
/db_xref="GI:4982402"  
/translation="MVLVDYGSQYSRLITRRIRENYSFVFPDDKVDLSKVDAVI  
LSGSPRSVIEDDAKLPEMFOEYKGPVLAICYGQQLVKELGSGVRRGREGYCTVE  
LSRPIFEGIEPEKHYMNSHODEVVRVLPGEHPIAVSETGVIATATGKRWLQIHP  
EVHHTTEYGDRIISNLFNVCKLEKNMKGIDLVKRIHRIKETIGNKRAIILASGVVS  
SVAALVHRAIGKMLVCVFVDHGLLRNREEREVRKEDNMVVDVDAKRKLELT  
RGVDPPEKRRITIGEEFLRVFEEBAKHDEYVQGRISDVIESAASGATTAKIKSH  
HNVGDLPEKMKLKIPEPLDLDFKDEVRKGVYIGIPRIINRHPFPGGLAVRYLGEV  
TEEKLEIREADYIFIEILRKHDYVDKVMGAFNLLPDKSVYKGDARAYEVVALVA  
VNSYGMTADRSRIPHDILBRARRITREYKGVGVYDITSRKPATIEVE"  
6282..6905  
/gene="TM1821"  
6282..6905  
/note="similar to GB:L77117 SP:Q58134 PID:1499541 percent  
identity: 72.33; identified by sequence similarity;  
putative"  
/codon_start=1  
/transl_table=11  
/product="conserved hypothetical protein"  
/protein_id="AAD36884.1"  
/db_xref="GI:4982403"  
/translation="MEELKLELERIREANPLVQRFEERKRLGEECTEEDLFCGLSE  
CVLVAANSABEGIAQAQKEIGEFYHLPLEEIAELRVRGHRYPQKRAEFTVENKRLG  
KLKMLVSGDPSREFPLVRNAKGIGMKEASHFLNTGVEDLAIIDKRVILMRKHGLI  
QEIKQSKSRKYLVEEILRKVAAPFSGSKFDLYLWLVKQKVDK"  
6917..7843  
/gene="TM1822"  
6917..7843  
/note="similar to GB:U09005 SP:P40605 PID:507734 percent  
identity: 58.45; identified by sequence similarity;  
putative"  
/codon_start=1  
/transl_table=11  
/product="f1sh protease activity modulator HflK"  
/protein_id="AAD36885.1"  
/db_xref="GI:4982404"  
/translation="MRKYVNVYVIVIGYFLNGYGVGSEVALKTFGRFYSVYS  
GIHYHLPIQSHVTVDTYTRKITEIFRSIQGERISQSPROEAMITGDNILVYS  
EAVQYVRKDPVAFAFNITADSIVRTTESVAREKXAMSIDVLTSGDEIGFKTA  
QMLQEIIDSYNCIGKVENYVLOEVVPPVDVADFDVNNAROKERLINSEARXYANDY  
VPRKOGAOEILROAEVAYAOEVYKALGEMKRFEEVLEYSKADITRRKMLDALQS  
LLEKSENKVFVFGDSLNLINISDLKMGK"  
7840..8691  
/gene="TM1823"  
7840..8691  
/note="similar to GB:U09005 SP:P40606 PID:507735 percent  
identity: 61.97; identified by sequence similarity;  
putative"  
/codon_start=1  
/transl_table=11  
/product="f1sh protease activity modulator HflC"  
/protein_id="AAD36886.1"  
/db_xref="GI:4982405"  
/translation="MKIWMISLLIILVNGAILLFSFFYVLDOTOQAVVRFKIVAY  
ETEPGLHFKQPFVDNVVRFDRILYIDIEKIIIAAKTLVLDITYLMIKIDAEAI  
KSLKSVKLAIDRIDDVYSHVRNIFAAANDEIISEREDLREVALISEADKQVFI  
EVDVVRVKAHLDPAENENKAYERRKARYSIAQIRAECEKARKIRAEADKTAQVLI  
AAQSKRKEQJIKGGEASAVKIYAEVFSKDDFEFWRTMEVYSIEKGIILIDDELDA  
LKYIKTK"  
complement(8704..9519)  
/gene="TM1824"  
complement(8704..9519)  
/gene="TM1824"  
/note="similar to percent identity: 0.00; identified by
```


DEFINITION	E.coli genes nusb (ssyB) and 3 orfs.
ACCESSION	X64395.1
VERSION	X64395.1 GI:42147
KEYWORDS	nusb (ssyB) gene.
SOURCE	Escherichia coli.
ORGANISM	Escherichia coli
REFERENCE	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
AUTHORS	Ito,K.
JOURNAL	1 (bases 1 to 2949)
TITLE	Direct Submission
REFERENCE	Submitted (06-FEB-1992) K. Ito, Inst for Virus Research, Kyoto University, Sakyo-ku, Kyoto 606-01, JAPAN
AUTHORS	2 (bases 1 to 2949)
TITLE	Taura,T., Ueguchi,C., Shiba,K. and Ito,K.
REFERENCE	Insertional disruption of the nusb (ssyB) gene leads to cold-sensitive growth of Escherichia coli and suppression of the secY24 mutation
JOURNAL	Mol. Gen. Genet. 234 (3), 429-432 (1992)
MEDLINE	93024316
FEATURES	
source	Location/Qualifiers
	1..2949
	/organism="Escherichia coli"
	/strain="K12"
	/db_xref="taxon:562"
	/map="9.5 min"
	330..357
	/note="promoter-like sequence P1"
	333..362
	/note="promoter-like sequence P1'"
	391..840
	/note="orf1"
	/codon_start=1
	/transl_table=11
	/protein_id="CAA45734.1"
	/db_xref="GI:42148"
	/db_xref="SWISS-PROT:P25538"
	/translation="MHCPGFADVTVISRLVGGSSVRRROCLGACNRFPTFEVA ELVMPRVKSNDRPEPNEKLSGMLRALEKRPVSSDYEMAINHKSQLRATGEVE VPSKMGVLMEQLKIDKRVAYIRFASVRSFEDIKFEGEIRLEAD"
	844..1947
	/note="orf2, homologue to B.subtilis ribG"
	/codon_start=1
	/transl_table=11
	/protein_id="CAA45735.1"
	/db_xref="GI:581147"
	/db_xref="SWISS-PROT:P25539"
	/translation="MODEYMARALKIAGRGFTTHPHNVGCVIVGDEIVGEGYHO RAEPRIEYHALRMAGEKAGATVYTLIEPSSHGRTPCCDALIAGVARYVASMDD PNPVGARGILRLQAGIDVSHGLMSEALQKKGFLKRRTRTEPTYLQKLGSLDR TAAVASESOWITSPQARRDVLHRAQSHALITTSATVYLADPALIVMSLDSQTL YPOONLROPRIIVIDSNNRTPVHRIVOQGETWFAFRTQDSSEWETVTLLIPEHK GHLDLVVLMQLOKQOINSIWEAGPTIAGALQAGLVDELIVYIAPKLIGSDARGIC TLGLEKLIADAPQFKREIRHVGPDVCLHVG"
	1957..1986
	/note="promoter-like sequence P2"
	2036..2506
	/note="orf3, homologue to B.subtilis ribH"
	/codon_start=1
	/transl_table=11
	/protein_id="CAA45736.1"
	/db_xref="GI:42150"
	/db_xref="SWISS-PROT:P25540"
	/translation="NNIIEANVATPDARVAITIRFNNFINDSLLEGADLGRIGOV KDNITIVVWPGAYELPELANGLAKTGKAYAVIALGTVIGGTAFHEFYVAGSNGILA HVMDSEIPIPAFCVLTTESIEQAIERAGTNAKKGAALTALEMIVLKAIA"
	2526..2945
	/gene="nusb (ssyB)"
	2526..2945
	/gene="nusb (ssyB)"
	/codon_start=1
	/transl_table=11
	/protein_id="CAA45737.1"

```
/db_xref="GI:581148"  
/db_xref="SWISS-PROT:P04381"  
/translation="MKPAARRARECAVOALYSWOLSONDIADVEYOFLABODYKVDV  
VLVRELLIAGVATNTKYLIDGLMKPYLSHLEELGOVEKAVLRILALYELSKRSDVPYKV  
AINEAIELAKSFQADSHKPYNGVIDLKAAPYIRPNK"  
BASE COUNT      708 a      739 c      825 g      677 t  
ORIGIN
```

```
Query Match      7.6%; Score 78.6; DB 2; Length 2949;  
Best Local Similarity 48.1%; Pred. No. 3.5e-05;  
Matches 290; Conservative 0; Mismatches 304; Indels 9; Gaps 2;  
  
QY      1 ATGATTATGCGATTCACATTGCAATAATGTCACAGTCAAGTCAAGTGTATATCCACC 60  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      862 ATGGGGCGGGCCCTTAAGCTGGCGCAGCAGAGACGTTTACCACGCATCCCAACCGAAT 921  
  
QY      61 GTTGGCGCTGTTGATTAATGAAGTAGAGATTGTTGTTGTTGTCACACTTGAGAAAA 120  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      922 GTGGGGTGGTCATGTCATAAGATGGCGAATTTGTGGTGAAGGTTACCACCAACGTGCG 981  
  
QY      121 GGTGACAGCATGCGGAGGTTCAAGCACTGATATGCGCACACAAATGCTGAAGTGG 180  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      982 GGTGACCAACATGCCAGTACACGCGTTGCGTATGGCGGGTGAATAAGCCAAAGGTGCG 1041  
  
QY      181 ACGATTATATACCTAGAGCCATGATGCTATTTGTTCAACACACCCTGTGTAAAC 240  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      1042 ACCGCTATGTACACTCGAACCCTGTAGCCATCATGTGCTAGCCACACCGTGTGAC 1101  
  
QY      241 AAAATTATGATTGTAAGTAGCAAAAGTAGTATACGCCAANAAGACATTCGTTAGAC 300  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      1102 GCACCTATCGCGCGTGGCGGTGGCGGCTTGCTCGATGCAAGATCCTAACCCGCGAG 1161  
  
QY      301 ACACATGGGATGAGACGTTACGGGCTCACGGTATTGAGTTGA-----ATGCGTTG 352  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      1162 GTCGCTGGCGCTGGACTTTACCGTCTGCACAGCGCTGGCATTGACGTACGCCACGGCCTG 1221  
  
QY      353 ATGATGAACGGGCATCACATTATACCAAGACTTTTAAAGCAAAAGCAAGCAACTGC 412  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      1222 ATGATGAGTGAAGCCGAGCAATGTAATAAGGCTTTCACAGCGATGGCGCGGCTTT 1281  
  
QY      413 CACAATTTACAGTGAAGTAT-CTGCAAGTTAGATGTTAAACAGCGAATGATATGGA 471  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      1282 CCTTATATTCAAGTTAAACTTGGCGCATGCTTGATGGTGCACGCGATGGCGAGCGGC 1341  
  
QY      472 CAAAGTCATGATTAACATAAAGGTTAAACAAGATGCTATAGTTAGTACATGGA 531  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      1342 GAAAGCCAGTGAATCACTTCGCCCCAGGCGCGCGTGAATGTAACACTACTGCGCGGCA 1401  
  
QY      532 CACGACGAGTGTTAAGTGAAGAGTACAGTTGAATTAGATGATCAATATATACTACA 591  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      1402 AGTCATGCCATTTTAACCAAGCAGCGCCACAGTGTGGCGGATGATCCTGCTTAACGGTG 1461  
  
QY      592 CGT 594  
      |||  
Db      1462 CGT 1464  
      |||
```

Search completed: November 19, 2000, 04:23:24
Job time: 12787 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 03:12:14 : Search time 118.16 Seconds
(without alignments)
3271.470 Million cell updates/sec

Title: US-08-978-456-1

Perfect score: 1029

Sequence: 1 ATGATTTGCGATTCACT.....TACTTACGAAGAAGTGA 1029

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_36.*
1: /SID56/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseqn/NA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1029	100.0	1029	20	Pyrimidine deaminase
2	1029	100.0	3336	20	Riboflavin-biosynt
3	1029	100.0	3336	20	S.-aureus-riboflav
4	1029	100.0	3336	20	Riboflavin-biosynt
5	1029	100.0	3336	20	S.-aureus-riboflav
6	1000	97.2	7588	18	Staphylococcus aur
7	352.2	34.2	1269	18	DNA encoding a Rib
8	352.2	34.2	1269	20	Pyrimidine deamina
9	120.4	11.7	5567	12	B.subtilis rib (ri
10	120.4	11.7	5567	20	B.subtilis rib op
11	120.4	11.7	5567	20	B.subtilis rib op
12	120.4	11.7	5567	20	Bacillus subtilis

13	120.4	11.7	5567	21	293752
14	114.2	11.1	14273	19	V52173
15	112.6	10.9	1101	20	X77179
16	112.6	10.9	3417	20	X77183
17	112.6	10.9	3417	20	X77574
18	112.6	10.9	3417	20	X79558
19	112.6	10.9	3417	20	V72454
20	77.6	7.5	546	19	Z6259
21	76.2	7.4	1230	20	X82297
22	76.2	7.4	4312	19	V27186
23	76.2	7.4	4312	20	X82299
24	54.4	5.3	1686	16	Q87587
25	51.2	5.0	3095	11	Q03875
26	47.8	4.6	9636	15	Q67190
27	47.2	4.6	1826	19	V29477
28	47.2	4.6	1826	19	V37413
29	46.6	4.5	6124	11	Q03568
30	46.4	4.5	5589	15	Q68195
31	44	4.3	1864	8	N71405
32	43.4	4.2	19124	18	T72882
33	43.4	4.2	19124	21	Z69287
34	43.2	4.2	933	20	X61528
35	43.2	4.2	1008	20	X61527
36	43.2	4.2	910715	20	X20248
37	43	4.2	1206	21	A27366
38	43	4.2	4248	21	A27364
39	42.8	4.2	162450	21	Z69657
40	42.6	4.1	2418	13	Q27886
41	42	4.1	3399	17	T05868
42	41.8	4.1	24601	20	X13160
43	41.6	4.0	4766	15	Q70102
44	41.2	4.0	1864	15	Q78892
45	41	4.0	58407	19	V21210

ALIGNMENTS

RESULT	1
X59909	
ID	X59909 standard; DNA; 1029 BP.
AC	X59909;
XX	
DT	04-AUG-1999 (first entry)
DE	Pyrimidine deaminase and pyrimidine reductase (ribG) DNA.
XX	
KW	Pyrimidine deaminase; pyrimidine reductase; ribG; antagonist;
KW	agonist; antimicrobial; antibiotic; Helicobacter pylori infection;
KW	H. pylori-induced cancer; gastrointestinal carcinoma; gastric ulcer;
KW	gastritis; ss.
OS	Staphylococcus aureus.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1029
FT	/tag="a
FT	/product="ribG"
XX	
PN	W09926475-A1.
XX	
PD	03-JUN-1999.
XX	
PF	23-NOV-1998; 98WO-US24857.
XX	
PR	25-NOV-1997; 97US-0978456.
XX	
PA	(SMK) SMITHKLINE BEECHAM CORP.
XX	
PI	Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DJC,
PI	Lonetto MA, Nicholas RO, Palmer LM, Pratt JM, Reichard RW,
PI	Rosenberg M, Tranl CM, Ward JM, Warren RL,

Rib operon of Bact
Streptococcus pneu
S. pneumoniae ribG
S. pneumoniae ribG
S. pneumoniae ribG
S. pneumoniae ribG
S. pneumoniae ribG
S. pneumoniae deri
App ribG gene sequ
Actinobacillus ple
Nucleotide sequenc
DNA encoding leuco
Sequence encoding
P. faiciparum tran
Orpinomyces cellu
Orpinomyces cellu
Sequence encoding
Corynebacterium ri
Sequence of ANS-1
Plasmodium var-7 g
Plasmodium var-7 p
B. burgdorferi ant
B. burgdorferi ant
Borrelia burgdorfe
Sinorhizobium meli
Sinorhizobium meli
Retinoblastoma bin
P.falciptarum GBP13
Chicken Leucocytoz
Enterococcus faeca
Malarial PfEMP3 ep
Aspergillus nidula
Methanococcus jann

XX WPI: 1999-347572/29.
DR P-PSDB: Y15913.
XX
XX
PT New Staphylococcus aureus polypeptide and polynucleotide useful in
PT the treatment of gastric ulcer and gastritis
XX
XX
PS Claim 6; Page 5; 48pp; English.
XX
CC The present sequence encodes a Staphylococcus aureus pyrimidine
CC deaminase and pyrimidine reductase (ribG) polypeptide. Staphylococcus
CC aureus ribG and its antagonists are used to treat individuals in
CC need of them. Disease related to expression or activity of ribG can
CC be determined by analysing the nucleic acid sequence encoding ribG
CC or detecting the ribG polypeptide in a sample. ribG can also be used
CC to identify antagonists or agonists. ribG, or its related nucleic acid
CC also has use as a vaccine to induce an immunological response in an
CC animal. Antimicrobial compounds (e.g. agonists and antagonists of ribG
CC especially broad-spectrum antibiotics, may be of use in the treatment
CC of Helicobacter pylori infection. This should decrease the advent of
CC H. pylori-induced cancers, such as gastrointestinal carcinoma. The
CC treatment should also cure gastric ulcers and gastritis.
SQ Sequence 1029 BP; 376 A; 162 C; 198 G; 293 T; 0 other;
XX

Query Match	100.0%;	Score 1029;	DB 20;	Length 1029;
Best Local Similarity	100.0%;	Pred. No. 3.6e-230;		
Matches 1029;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

OY	1	ATGATTTATGGATTCACACTTGCAAATNTGGTACAAAGTCAAAAGGTGTTATCCACC	60
Db	1	atggattatgcatccaactctgcaaatatggtataaaggtcaaaaggtgttaataccacc	60
OY	61	GTTGCGCGTGTGTGATTAATGAAGGTGAGATTGGTATTGGTGCACCTTGAGAAA	120
Db	61	gttgcgcgtgtgtgtatgaatgaaggtatggtgttattggtgcacacttgagaaa	120
OY	121	GGTGCAAGCATGCGGAGGTTCCAAAGCTGATATGCGACAAACAAATGCTGAAGTGG	180
Db	121	ggtgcaagcatgcygagagttcaagcactgatatgycacacaaaatgctgaagtgcg	180
OY	181	ACGATTTATATTAGCTTGAGAGCAATGTCATATTGGTTCAACACACCCGTGTTAAC	240
Db	181	acgatttatatacgtatagaccatgtagtcattcgtgtcaacacacccctgtttaac	240
OY	241	AAATTTATGATTGTGAAGATGACAAAGTAGTATACGCACAAAGAACAAATTGCTAGAC	300
Db	241	aaattattgtattgtatagatagcaaanagtagtatacgcacaagaacattcttagac	300
OY	301	ACACATGGTGATGAGACCTTACGGGCTACGGTATTGAGTTGAATGCGTTGATGATGAA	360
Db	301	acacatggtgatgagacctacggctacggctacggtattgagttgaaatgcgttgaatgaa	360
OY	361	CGGCGATGCACAATTATTACCAAGACTTTTTTTAAACCAAAAGCAAGCACTGCCAATT	420
Db	361	cggcgatccaacattataccaagactctttttaagcaaaagcaagcaactgcgccaatt	420
OY	421	ACAGTGAAGATATCTGCAAGTTTAGATGTTAACAACGCAATGATTAATGCAACAAATCAA	480
Db	421	acagtgaaagatctcgcgaagtttagtgttaacaaggaatgataatgyscaagfcaa	480
OY	481	TGGATTACTTAACAAAGAGGTTAAACAAGATGCTATTAAGTTAAACATCMGACACGAGCA	540
Db	481	tggattacttaacaagaggtttaaaaagatgctataagttlaagacatcgcaagagca	540
OY	541	GTTGTTAACTGAGACGCTACAGTTGAATTAATGATCCACAATATTACTACAGATTCAA	600
Db	541	gttctaactggaagcgtacacggttgaattlaagatgccacatatctactacgfatcaa	600
OY	601	GATGAAAAAAACCTATTAAGATAATTGTCTAAGTCTGGGAATATTCAATTTTAATCAG	660
Db	601	gatgaaaaaacccataaaagtaatatctgtctaagtcctggaatatactatitaaacag	660

QY	661	CAATTATTCAGATGCATACACCAATTGGATATATACGAAATCCAAATTTAA	720
Db	661	caatttccaagtcgatataacaccaatttggatatactcgaataatccaatttaca	720
QY	721	AGCATCAACACATTTGAATTTATTTACTGTTAAATCTTGATTTAACAACTGTT	780
Db	721	agcaatcaaacacatattgaattatattacttgaagtccttggatttaacaaacttct	780
QY	781	CACATTTATATAAAGAGGAGTTGGAACTTTGCTAGTGGAGGACAGTCCAACTACT	840
Db	781	cacaattatataaagaggaagtcgacttctagtcgagcgaggtccaacactact	840
QY	841	TCAGATTTCTCCATCTTTATATAGATGAATTTATCTCTATATATGCCCGAAATTA	900
Db	841	tcagaattctccactctatataatagaatgaattatctctataatgccccgaataat	900
QY	901	GCGGATCTGGAATTTTCATTTTTCACAAACAATGATGTATTGAGTACACAGTCG	960
Db	901	gcggaatctggaattctcaatttatcaaacacaaatgatatgagataccgaatcg	960
QY	961	AACCAATTTGAATTTGTTTCATTCGCGATTATTAATCAAAATGTTAAATTAATTA	1020
Db	961	aaccaatttgaattgttcattccgcggtatctaataacaaatglttaattacttcaga	1020
QY	1021	AAGAAAGTGA 1029	
Db	1021	aaagaagtga 1029	

RESULT 2

ID	X59913	standard; DNA; 3336 BP.
----	--------	-------------------------

AC X59913;

DT 04-AUG-1999 (first entry)

DE Riboflavin biosynthesis operon containing the ribG gene.

KW pyrimidine deaminase; pyrimidine reductase; ribG; antagonist;

KW H. pylori-induced cancer; gastrointestinal carcinoma; gastric ulcer;

KW gastritis; ss.

OS Staphylococcus aureus.

PN W099926475-A1.

PD 03-JUN-1999

PF 23-NOV-1998; 98WO-US24857.

PR 25-NOV-1997; 97US-0978456.

PA (SMIK) SMITHKLINE BEECHAM CORP.

PI Black MT, Burnham MKR, Fedon J

PI Rosenberg M, Trani CM, Ward JM, Warren RL;
.....

DR WPI: 1999-347572/29.

PT New Staphylococcus aureus polypeptide and polynucleotide useful in

XX

XX

CC riboflavin biosynthesis operon containing the pyrimidine

CC aureus ribg and its antagonists are used to treat individuals in

[illegible]

```

OY 1021 AAGAGTGA 1029
    |||||
Db 1021 aagaagtga 1029

RESULT 5
X60830
ID X60830 standard; DNA: 3336 BP.
AC X60830;
XX
XX 09-AUG-1999 (first entry)
DE
S. aureus riboflavin biosynthesis operon.
XX
XX Lumazine synthase; ribH; riboflavin; differential screening; gastritis;
XX H. pylori infection; gastrointestinal carcinoma; gastric ulcer; wound;
XX bacterial adhesion; cell invasion; riboflavin biosynthesis operon; ss.
XX
XX Staphylococcus aureus.
XX
XX W09927128-A1.
XX
XX 03-JUN-1999.
XX
XX 24-NOV-1998; 98WO-US25096.
XX
XX 25-NOV-1997; 97US-0977553.
XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX
XX Debouck C, Fedon JC, Jaworski DD, Mooney J, Palmer LM;
XX Traini CM, Wang M, Warren RL, Zhong Y;
XX
XX WPI: 1999-357850/30.
XX
XX New Staphylococcus aureus ribH polynucleotide and polypeptide useful
XX in the treatment of gastric ulcers and gastritis
XX
XX
XX Example 2; Page 34-36; 48pp; English.
XX
XX The invention relates to polynucleotides and polypeptides of the lumazine
XX synthase (beta -subunit) family, designated ribH. They are isolated from
XX Staphylococcus aureus. ribH enzymes are involved in the synthesis of
XX riboflavin. The polynucleotides and polypeptides may be employed as
XX research reagents and material for the discovery of treatments and
XX diagnostics for diseases, particularly human diseases. They can be used
XX for diagnosis of the disease and staging of disease, and as reagents in
XX differential screening methods. The polynucleotides may be used as a
XX source for hybridization probes, and for screening of genetic mutations,
XX serotype, and identification, and for chromosome identification. The
XX polypeptides can be used to produce antibodies. The polypeptides can
XX also be used in vaccine formulations, and to identify agonists and
XX antagonists. These are used to prevent, inhibit or treat diseases,
XX particularly of Helicobacter pylori infections, such as gastrointestinal
XX carcinoma. Gastric ulcers and gastritis may also be treated. The
XX polypeptides can also be used to treat wounds and in-dwelling devices to
XX prevent bacterial adhesion and infection, and to block ribH protein-
XX mediated mammalian cell invasion. Antagonists and agonists of ribH may
XX also be employed to inhibit and treat diseases. The present sequence
XX represents the S. aureus riboflavin biosynthesis operon that encodes
XX genes ribG, ribB, and ribH.
XX
XX Sequence 3336 BP; 1202 A; 459 C; 678 G; 997 T; 0 other;
XX
XX
XX Query Match 100.0%; Score 1029; DB 20; Length 3336;
XX Best Local Similarity 100.0%; Prid. No. 4.9e-230;
XX Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 ATGGAATTATGGCATTCACACTTGCAATATGTCATCAAGGCAACGATGTTATATCAACC 60
XX |||||||

```

```

Db      1 atgatatatgcattcaacttcgcaaatatgtaacaagtcgaacagtgtaatccacc 60
Oy      61 GTTGGCGCTGTTGTATGAAGTAGAGTTGTTGATTTGGTGCACACTTGAGAAA 120
Db      61 gttggcgctgtgtatgaagtaagtagatctgttcattgtgcacaacttgaga 120
Oy     121 GGTGACAGCATGCGAGGTTCAACACTTGTATATGACCAACAATAATGCTGAGTGG 180
Db     121 gttgcaagcatgcgagggttcaagcacttgatgacgcaacaataatgctgaagtg 180
Oy     181 ACGATTATATTACTTATAGCCATGTACTCAATTTGGTTCAACACCCTGTTTAAC 240
Db     181 acgattatattacgttagagccatgtatcttgtgttcacaccccccgtgttaac 240
Oy     241 AAAATTATGATTGTATAGTATACCAAGATATACGCAACAAAAGCAATTGTTAGC 300
Db     241 aaaattatgtatcgaatgaacaaagatagatcgcacaacaaagacatcgttagac 300
Oy     301 ACACATGCTGATGAGACGTTTACGGGCTCAGCGTATTTGAAGTTGAATGCGTTGATGAA 360
Db     301 acacatgtgtatgagagcgtttagcggtcagcggtattgaggttgatgagaa 360
Oy     361 CGGCATCACAAATTATACCAAGACTTTTAAAGCAAAAAGCAAGCACTGCCACAATT 420
Db     361 cggcatcacaaattataccaagactttttaaagcaaaagcaaacactgcacaact 420
Oy     421 ACAGTGAAGTATCTGCAAGTTAGTGTAAACAAGCAATGATATGACAAAGTCAA 480
Db     421 acagtgaagatctcgaagtttagatgtaacaagcaagcaaaatgaatgaagaagtc 480
Oy     481 TGGATTACTTAACAAAGAGTTTAAACAGATGCTATTAAGTTAAGACATGACACGACA 540
Db     481 tggattacttaacaaagaggttaacaagatgtctaaagttgaagcatcgcacagca 540
Oy     541 GTGTTAACTGGAAGCGTACACTTGAATTAGATGATCCACATATACTACACTTATCAA 600
Db     541 gtgttaactggaagcgtttagatgtaagttgaatgataccacaatactacacgtatt 600
Oy     601 GATGAAAAAACCTATATAAAGTATATGCTAAGTGGGAAATATTCATTTTATATCG 660
Db     601 gatgaaaaaacctataaagtaaatatgcttaagtcgggaatatcattcaatccag 660
Oy     661 CAATTTATCAAGATGAATCAACACCAATTTGATATATATCTGAAATTCAAATTTACA 720
Db     661 caattatcaagatgaatcaacaccaatttgatataatacgaataatccaaattaca 720
Oy     721 AGCAATCAACACATATTTGAATTTATTACTTGAAGCTGTGATTTTAAACAATTTCT 780
Db     721 agcaatcaaacacatatgaattatttactgaagctgtgatttacaacaattctt 780
Oy     781 CACATTTATATAAAGAGGAGTTGAACTTGCTAGTGAAGCAGGTCACCACTACT 840
Db     781 cacatTTATATAAAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 840
Oy     841 TCAGATTCTCCATCTATTATATAGATGAATTTCTCTATATATGCCCGAAATTAAT 900
Db     841 tcagattctccatctattatataagatgaattattctctattatagcccggaataatt 900
Oy     901 GCGGATCTGCAAAATTTATCAATTTTATCAACAAATGATGATTTGAGATACCAAGTGG 960
Db     901 gcggatctgcaaaatTTATCAATTTTATCAACAAATGATGATTTGAGATACCAAGTGG 960
Oy     961 AACCAATTTGAATTTGTTTATCCGAGTTATATCAAAATGTTAAATTAATTAATTAACA 1020
Db     961 aaccatttgaatattgttcatccgagttatcaaatcaaatgtaataacttaacttaca 1020
Oy    1021 AAGAAGTGA 1029
Db    1021 aagaagtga 1029

```

RESULT 6
V74453/c

```

ID      V74453 standard; DNA; 7588 BP.
XX
AC      V74453;
XX
DT      16-MAR-1999 (first entry)
XX
DE      Staphylococcus aureus config SEQ ID #142.
XX
KW      Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW      cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW      skin infection; surgical wound infection; scalded skin syndrome;
KW      toxic shock syndrome; ds.
XX
OS      Staphylococcus aureus.
XX
FH      Key Location/Qualifiers
FT      misc-feature 651..720
FT      /*tag= a
FT      /note= "these bases represent a line of missing text in
FT      the sequence listing in the specification. They
FT      are included to maintain the nucleotide numbering
FT      given in the specification for this DNA sequence"
FT      misc-feature 2461..2520
FT      /*tag= b
FT      /note= "these bases represent a line of missing text in
FT      the sequence listing in the specification. They
FT      are included to maintain the nucleotide numbering
FT      given in the specification for this DNA sequence"
FT      misc-feature 4261..4320
FT      /*tag= c
FT      /note= "these bases represent a line of missing text in
FT      the sequence listing in the specification. They
FT      are included to maintain the nucleotide numbering
FT      given in the specification for this DNA sequence"
FT      misc-feature 6061..6120
FT      /*tag= d
FT      /note= "these bases represent a line of missing text in
FT      the sequence listing in the specification. They
FT      are included to maintain the nucleotide numbering
FT      given in the specification for this DNA sequence"
XX
PN      EP786519-A2.
XX
PD      30-JUL-1997.
XX
PE      07-JAN-1997; 97EP-0100117.
PR      05-JAN-1996; 96US-0009861.
XX
PA      (HUMA-) HUMAN GENOME SCI INC.
XX
PI      Barash SC, Chol GH, Dillon PJ, Fannon MR, Kunsch CA;
PI      Rosen CA;
XX
XX      WPI; 1997-374922/35.
XX
DR      Polynucleotide(s) and proteins derived from Staphylococcus aureus
XX      stored on computer readable medium and used in the production of
XX      anti-S.aureus vaccines
XX
XX      Claim 1; Page 740-744; 3271pp; English.
XX
XX      This sequence represents one of 5191 Staphylococcus aureus DNA sequences
XX      of the invention. The DNA sequences are recorded on a computer readable
XX      medium, preferably selected from a floppy or hard disk, random access
XX      memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
XX      the S.aureus DNA sequences allows putative functions to be assigned so
XX      that protein-encoding or regulatory regions of commercial, therapeutic or
XX      industrial importance can be obtained. Specifically, sequences which are
XX      likely to encode antigens have been identified and these polypeptides can
XX      be used in a vaccine composition against S.aureus infection. The
XX      polypeptides can also be used in a kit for the immunodetection of
XX      S.aureus in a sample. S.aureus is implicated in numerous human diseases,

```

including cellulitis, eyelid infections, food poisoning, osteomyelitis, skin and surgical wound infections, scaled skin syndrome, toxic shock syndrome, etc. Organisms transformed with the DNA sequences can be used for recombinant production of the polypeptides. The new DNA sequences (and their fragments) are useful as primers or probes for isolating homologues of any of the *S. aureus* DNA sequences contained on the computer readable medium.

Sequence 7588 BP; 2333 A; 1341 C; 1061 G; 2613 T; 240 other:

Query Match 97.2%; Score 1000; DB 18; Length 7588;
Best Local Similarity 98.7%; Pred. No. 3,3e-223;
Matches 1019; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

```

OY 1  ATGGATTATGCGATTCACCTTGCAAAATATGGTACAAAGTCAAAAGGTGTTAATCCACC 60
    |||||||
DB 5849 ATGGATTATGCGATTCACCTTGCAAAATATGGTACAAAGTCAAAAGGTGTTAATCCACC 5790
OY 61  GTTGGGGCTGTTAGTATGAAGTATGTTGGTGTGGTGCACCTTGAGAAA 120
    |||||||
DB 5789 GTTGGGGCTGTTAGTATGAAGTATGTTGGTGTGGTGCACCTTGAGAAA 5730
OY 121 GGTGACACGATCGGAGGTTCAAGCACTGATATGGACACAAAGTCTGAAGTGC 180
    |||||||
DB 5729 GGTGACACGATCGGAGGTTCAAGCACTGATATGGACACAAAGTCTGAAGTGC 5670
OY 181 ACGATTATATTCAGTTAGAGCCATGATGATTTGGTTCAACACCCCTGTGTTAC 240
    |||||||
DB 5669 ACGATTATATTCAGTTAGAGCCATGATGATTTGGTTCAACACCCCTGTGTTAC 5610
OY 241 AAAATTATGATTTGAAGATAGCAAAAAGTATAGCAACAAAGCATTTGTTAGAC 300
    |||||||
DB 5609 AAAATTATGATTTGAAGATAGCAAAAAGTATAGCAACAAAGCATTTGTTAGAC 5550
OY 301 ACACATGATGATGAGAGCTAGCGGCTCAAGTATGAGGTTGAATGCGTTGATGAA 360
    |||||||
DB 5549 ACACATGATGATGAGAGCTAGCGGCTCAAGTATGAGGTTGAATGCGTTGATGAA 5490
OY 361 CGGCGATCAACAATTATACCAAGACTTTTAAAGCAAAAGCAAAAGCAACTGCACAAAT 420
    |||||||
DB 5489 CGGCGATCAACAATTATACCAAGACTTTTAAAGCAAAAGCAAAAGCAACTGCACAAAT 5430
OY 421 ACGTGAAGTATCTGCAATTTAGATGTTAAACAAGCCGATGATATGACAAAGTCAA 480
    |||||||
DB 5429 ACGTGAAGTATCTGCAATTTAGATGTTAAACAAGCCGATGATATGACAAAGTCAA 5370
OY 481 TGGATTACTACAAAGGTTAAACAAGATGCTATAGTTAAGACATCGACACGACGA 540
    |||||||
DB 5369 TGGATTACTACAAAGGTTAACAAGATGCTATAGTTAAGACATCGACACGACGA 5310
OY 541 GTGTTACTGGAAGAGCTACAGTTGAATTAGATGATCCACAATATACACGATTTCAA 600
    |||||||
DB 5309 GTGTTACTGGAAGAGCTACAGTTGAATTAGATGATCCACAATATACACGATTTCAA 5250
OY 601 GATGAAAAAACCCCTATAAAGTAATTTGTCTAAGTCTGGGAATATCTCTTTTAATCAG 660
    |||||||
DB 5249 GATGAAAAAACCCCTATAAAGTAATTTGTCTAAGTCTGGGAATATCTCTTTTAATCAG 5190
OY 661 CAAATTATCAAGATGAATACACCAATTTGATATATCTGAAAAATCCAAATTTTACA 720
    |||||||
DB 5189 CAAATTATCAAGATGAATACACCAATTTGATATATCTGAAAAATCCAAATTTTACA 5130
OY 721 AGCAATCAACACATATTTGAATATTTACTTGAAGTCTTGATATTTAACACAATCTT 780
    |||||||
DB 5129 AGCAATCAACACATATTTGAATATTTACTTGAAGTCTTGATATTTAACACAATCTT 5070
OY 781 CACAATTTATATAAAGAGAGTGTGACCTTGTCTAGTCGACGAGGTCCACCACTACT 840
    |||||||
DB 5069 CACAATTTATATAAAGAGAGTGTGACCTTGTCTAGTCGACGAGGTCCACCACTACT 5010
OY 841 TCGAATTCCTCCATCT--TTATATAGATGAATTTATCTCTATTTATGCCCGGAATTA 897
    |||||||

```

```

DB 5009 TCGAATTTCTCCAAATCTAATTATATAGATGAATTTATCTGATATGCCCGGAATTA 4950
OY 898 ATGGGGGATCTGGAATATATCAATTTTATCAACAAATGATGTGATGATACAGAT 957
    |||||||
DB 4949 ATGGGGGATCTGGAATATATCAATTTTATCAACAAATGATGTGATGATACAGAT 4890
OY 958 GCGAACCAATTTGAATTTGTCATTCGAGTTATTTAAATCAAAATGTTAAATTAACCTTA 1017
    |||||||
DB 4889 GCGAACCAATTTGAATTTGTCATTCGAGTTATTTAAATCAAAATGTTAAATTAACCTTA 4830
OY 1018 CGAAGAGAGTGA 1029
    |||||||
DB 4829 CGAAGAGAGTGA 4818

```

```

RESULT 7
ID T84023
T84023 standard; DNA; 1269 BP.
XX
AC T84023:
XX
DT 01-SEP-1998 (first entry)
XX
DE DNA encoding a Riboflavin biosynthesis protein RIB6.
XX
KW Staphylococcus aureus protein; ribozyme; antisense sequence; control;
KW Staphylococcal gene; regulatory element; bacterial gene expression;
KW vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;
KW toxic shock syndrome; ss.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT CDS 734..1033
FT /tag= a
XX
PN MO9730070-A1.
XX
PD 21-AUG-1997.
XX
PE 19-FEB-1997; 97WO-US02318.
XX
PF 20-FEB-1996; 96US-0011888.
PR (SMK ) SMITHKLINE BEECHAM CORP.
XX
PA Black MT, Burnham MK, Hodgson JE, Knowles DUC, Nicholas RO;
PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;
PI WPI: 1997-424969/39.
DR F-PSDB: W28084.
DR
XX
PT Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used
PT to isolate antimicrobial compounds, and in vaccines against S.
PT aureus infection
XX
PS Claim 9; Page 825; 989pp; English.
XX
CC The present sequence encodes a Staphylococcus aureus protein, that,
CC based on homology with an Bacillus subtilis protein, is believed to be a
CC Riboflavin biosynthesis protein RIB6. The present sequence was obtained
CC from a library of clones of S. aureus WCUH 29 in Escherichia coli. The
CC DNA sequence can be used in the construction of ribozymes and antisense
CC sequences to control the expression of Staphylococcal genes. The DNA
CC sequence is also useful as a source of regulatory elements for the
CC control of bacterial gene expression. The encoded protein may be used
CC to produce vaccines to enable a host to produce specific antibodies
CC with antibacterial action. These vaccines and antibodies would protect
CC a host against invasion by S. aureus, and conditions relating to
CC Staphylococcal infection, e.g. Staphylococcal food poisoning, scaled
CC skin syndrome, and toxic shock syndrome.
XX
SQ Sequence 1269 BP; 435 A; 188 C; 256 G; 371 T; 19 other:

```


XX	B. subtilis rib operon nucleotide sequence.
DE	Riboflavin; open reading frame; ORF; structural gene; promoter;
XX	Vitamin B2; Bacillus subtilis; rib operon; ds.
KW	
XV	
XX	
OS	Bacillus subtilis.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	1..5567
FT	/tag= a
FT	/note= "reading frame 1 translated protein (Y21801)"
FT	2..5567
FT	/tag= b
FT	/note= "reading frame 2 translated protein (Y21802)"
FT	3..5567
FT	/tag= c
FT	/note= "reading frame 3 translated protein (Y21803)"
XX	
PN	US5925538-A.
XX	
PD	20-JUL-1999.
XX	
PF	24-AUG-1998; 98US-0138775.
XX	
PR	11-SEP-1990; 90US-0581048.
PR	22-JUN-1989; 89US-0370378.
PR	21-APR-1992; 92US-0873572.
PR	06-FEB-1995; 95US-0384626.
PR	24-AUG-1998; 98US-0138775.
XX	
PA	(HOFF) ROCHE VITAMINS INC.
PI	Erdenberger T, Hatch RT, Hermann T, Perkins JB;
PI	Peto JG, Sloma A;
XX	
DR	WPI; 1999-418271/35.
DR	P-PSDB; Y21801, Y21802, Y21803.
XX	
PT	Use of bacterial strains for the over production of riboflavin
XX	
PS	Example 2; Fig 3A-S; 61pp; English.
XX	
CC	The invention relates to a method of producing riboflavin that comprises,
CC	providing a recombinant bacterium containing an exogenously introduced
CC	bacterial ORF (open reading frame) 3 (or ORF5) structural gene and an
CC	exogenous promoter, culturing the bacterium, and recovery of the product.
CC	The method is useful for the production of large quantities of riboflavin
CC	(vitamin B2). The present sequence represents the complete nucleotide
CC	sequence of B. subtilis rib operon (also see X81948 for the antisense
CC	sequence).
XX	
SO	Sequence 5567 BP; 1751 A; 1096 C; 1297 G; 1423 T; 0 other:
XX	
Query Match	11.7%; Score 120.4; DB 20; Length 5567;
Best Local Similarity	51.0%; Pred. No. 4,9e-19;
Matches 317; Conservative	0; Mismatches 296; Indels 9; Gaps 1.
Dy	1 ATGATTATTCGACCTTCAACTGCMAATATGTACAAAGTGTTAATCCACCC 60
Db	1116 attgaagcggcgcccttgatcttcgaaagcggcgaaagcacgcgaatccatccgcgc 1175
Oy	61 GTTGCGCGCTGTTGTAAGTAATGAAGGTAGATTTGGTATTGGTCACACTTGAGAAAA 120
Ob	1176 gtcggcgctgttgcgtlaagaagcagacaattgtcgsagtggcgcccatttaaatat 1235
Oy	121 GGTCACAAGCATGCCGAGGTTCAAGCAGCTGTATGGCACACAAAATCTGAAGTGGC 180
Db	1236 gttgaagctcaaggcaaaattcatccatcataatgctggagacatatcagaagggtgcc 1295
Oy	181 AGGATTATATACGTTAACGCCAGTAGTACATTTTGGTTCAACACCACCCGTGTATAC 240

Dd	1296	gaatttaacgtlaaccctcgcaacccgctcattcaagsgaaaaaaacaaccgcctatgttcgaga	1355S
Oy	241	AAATATTGATTGTTAAGATACGAATAAAGTATGTTACGCACAATAAGA-----CAAT	291
Dd	1356	ttagattcaacctcttgatccaagaagatgttctgtgcgatbagatctcatccgctt	1415S
Oy	292	TGGTTAGCACACATCGTGATGAGACGTTTCACGGCTCACCAGTATTGAGGTTGAATGCCGTT	351
Dd	1416	gtgcctgnaaaggatcacgatgatgaaaagaagctggcattgaggttaaggyaaagcatc	1475S
Oy	352	GATGATGACCCGGCATCCACTATTATACCAAGACTTTTTTTAAAGCAAAAAGCAAGCAACTG	411
Dd	1476	ctggagagcccgagggagagcgctgaatgaaaaattctgcactttatgagcgagcgctt	1535S
Oy	412	CCACAATTACAGTGAAGATATCTGCAAGTTTATGATGTGAACCAAGCGATATATATGA	471
Dd	1536	cctgactgcacgcttaaaagcgctgccagccttgacgysaaatgatcaccagccagcgggt	1595S
Oy	472	CAAAGTCATGATGATTACTAACAAGAAGGTTAAACAAGATGCTATTAAGTTAAGACATCGA	531
Dd	1586	gaccagcaaatggataccagcttaagagcgctgcagaacagatgctccagcaatacaggaaca	1655S
Oy	532	CACGACGCAGTGTACTGGAAGAACGATACAGTTGAAATTGATGATCCACATATACTACA	591
Dd	1656	caccaagaacttattagtcggagtgtgcacagtgaaagccgacaaatccgagcttaaccgc	1715S
Oy	592	CGTATTCAAGATGCAAAAAAC	613
Dd	1716	agactgccgaatgtaacaaaac	1737
 RESULT_11 X81948/c ID X81948 standard; DNA; 5567 BP. XX XX X81948; AC XX DT 10-SEP-1999 (first entry) XX DE B. subtilis rib operon complementary nucleotide sequence. XX KW Riboflavin; open reading frame; ORF; structural gene; promoter; KM vitamin B2; Bacillus subtilis; rib operon; complementary; ds. XX OS Bacillus subtilis. XX FH Key Location/Qualifiers FT CDS 1..5567 FT /tag= a FT /note= "reading frame 1 translated protein (Y21804)" FT CDS 3..5567 FT /tag= b FT /note= "reading frame 3 translated protein (Y21805)" FT CDS 2..5567 FT /tag= c FT /note= "reading frame 2 translated protein (Y21806)" XX XX US592538-A. PN PD 20-JUL-1999. XX XX 24-AUG-1998; 98US-0138775. XX PR 11-SEP-1990; 90US-0581048. PR 22-JUN-1989; 89US-0370378. PR 21-APR-1992; 92US-0873572. PR 06-FEB-1995; 92US-0384626. PR 24-AUG-1998; 98US-0138775. XX XX (HOFF) ROCHE VITAMINS INC. PI Erdemberger T, Hatch RT, Hermann T, Perkins JB; PI Pero JG, Sloma A;			

XX MPI: 1999-418271/35.
DR P-PSDB: Y21804, Y21805, Y21806.
XX
PT Use of bacterial strains for the over production of riboflavin
XX
PS Example 2; Fig 3A-S; 61pp; English.
XX
CC The invention relates to a method of producing riboflavin that comprises,
CC providing a recombinant bacterium containing an exogenously introduced
CC bacterial ORF (open reading frame) 3 (or ORF3) structural gene and an
CC exogenous promoter, culturing the bacterium, and recovery of the product.
CC The method is useful for the production of large quantities of riboflavin
CC (vitamin B2). The present sequence represents the complementary
CC nucleotide sequence of B. subtilis rib operon.
XX
SQ Sequence 5567 BP; 1422 A; 1296 C; 1097 G; 1752 T; 0 other;

Query Match 11.7%; Score 120.4; DB 20; Length 5567;
Best local Similarity 51.0%; Pred. No. 4,9e-19;
Matches 317; Conservative 0; Mismatches 296; Indels 9; Gaps 1;

QY 1 ATGATTATGCGATTCACTTGCATAATATGCTACAGGTCAACAGGTGTTATCCACC 60
DB 4452 ATGAGCTGGCCTTAGATCTTGCAGACAGGCGGAGAGACAGACCAATCCATCCGCTC 4393

QY 61 GTTGGCCCTGTGTAGTTAATGAAGTAGAGATTGTTGTTATGTTGTCACACTTGAAGAAA 120
DB 4392 GTGGCCCTGTGTGTAAAGGAGGACCAAAATGTCGAAATGGCGCCATTAAATAT 4333

QY 121 GGTGACAGCATGGCGAGGTTCAGACACTGATGTGACACAAATGCTGAAGGTGCG 180
DB 4332 GGTGACACTATGACAGATTTCATCCATATGCTGTGAGACATGCGAGGAGTGCC 4273

QY 181 ACGATTATATAGCTTAGAGCCATGTAGTATTTGGTTCAACACCACTGTGTAC 240
DB 4272 GACATTATACCTTACACTGCAACCGTCGACCATACGAAACACCGCATGTGAGAA 4213

QY 241 AAAATTTTGATGTAGATAGCAAAAGTAGTATACCCACAAAGA-----CAAT 291
DB 4212 TTGATTATCAACTCTGTATCAAAAGAGTGTCTGCGATGAGAGATCCATATCCGCTT 4153

QY 292 TCGTTACACACATGTGTATGAGCGTTACGGGCTACAGGTATGAGTTGATGCGTT 351
DB 4152 GTGGCTGGAAGAGGATCAGCATGATGAAGACGTGCAATTGAGTAAGGAAAGSCATC 4093

QY 352 GATGATGACGGGATCAACAATTATACCAAGACTTTTAAAGCAAAAGCAAGCAACTG 411
DB 4092 CTGGCAGACAGCGGAGAGGCTGAATGAATAATTTCTGCACCTTATAGAGACAGGCTT 4033

QY 412 CCACAAATTTACAGTAAGATCTGCAAGTTTGTAGATGTAAACAGGAAATGATATGA 471
DB 4032 CCGTACCTGCTCAAAAGCGGCTGCCAGCTTGACGGCAAGATACCAAGCAGCGGT 3973

QY 472 CAAGTCAATGAGATTACTAACAAGAGTTAAACAAGATGCTATAGTTAAGCATGA 531
DB 3972 GACGCAAAATGATCAGTCAGAGCTGCAAGAGAGATGCTCGCATACAGGAAACA 3913

QY 532 CACGACGCACTGTTAACTGAGACGTAAGTGAATTAATGATGCACATATACTACA 591
DB 3912 CACCAAGCATTTTAGTCGGAGTTGGCAGCATGTAAGAACCGACATCCGAGCTTAACCTGC 3853

QY 592 CGTATTCAAGATGGAACCAAC 613
DB 3852 AGACTGCCGATGTAAACAAAC 3831

RESULT 12
V71600
ID V71600 standard: DNA; 5567 BP.
XX
AC V71600;

XX 01-MAR-1999 (first entry)
DT
XX Bacillus subtilis rib operon.
XX DE
XX Riboflavin; vitamin B2; riboflavin synthase; deaminase; reductase;
KW GTP cyclohydrolase; rib operon; ds.
XX
XX Bacillus subtilis strain 168.
OS
XX Key Location/Qualifiers
FH RBS 352..357
FT RBS /tag= a
FT CDS 364..681
FT /tag= b
FT /label= ORF6
FT 708..749
FT /tag= c
FT /note= "rho-independent transcription terminator"
FT 771..799
FT /tag= d
FT /label= PI
FT 1034..67
FT terminator
FT /tag= e
FT /note= "rho-independent transcription terminator"
FT 1184..1195
FT /tag= f
FT 1422..1433
FT /tag= g
FT 1437..2186
FT CDS /tag= h
FT /label= ORF5
FT 2178..2194
FT /tag= i
FT 2197..2844
FT /tag= j
FT /label= ORF4
FT 2528..2556
FT /tag= k
FT /label= P2
FT 2843..2854
FT /tag= l
FT 2859..4055
FT /tag= m
FT /label= ORF3
FT 4071..4085
FT /tag= n
FT 4088..4552
FT /tag= o
FT /product= riboflavin synthase beta subunit
FT 4545..4574
FT /tag= p
FT /label= P3
FT 4653..4659
FT /tag= q
FT 4665..5039
FT /tag= r
FT /label= ORF2
FT 5038..5090
FT terminator
FT /tag= s
FT /note= "rho-independent transcription terminator"
FT complement (5053..5567)
FT /tag= t
FT /label= ORF1

US5837528-A.
XX
XX 17-NOV-1998.
XX
XX 06-FEB-1995; 95US-0384626.
XX 11-SEP-1990; 90US-0581048.
XX 22-JUN-1989; 89US-0370378.
PR

PR 21-APR-1992; 9205-0873572.
 PR 06-FEB-1995; 9505-0384626.
 XX
 PA (HOFF) HOFFMANN LA ROCHE INC.
 PI Perkins JB, Pero JG, Sloma A;
 DR WPI: 1999-023456/02.
 DR P-PSDB: W81549-55.

PT Nucleic acid encoding proteins involved in bacterial synthesis of
 PT riboflavin - also mutant or recombinant bacteria producing
 PT riboflavin at high levels
 XX
 PS

Example 7; Fig 3A-S; 62pp; English.

CC This is the nucleotide sequence of the rib operon of *Bacillus*
 CC *subtilis*. The rib operon includes an open reading frame (ORF)
 CC encoding the beta-subunit (see W81553) of riboflavin synthase, and
 CC 6 other ORFs, of which ORF1 is a partial coding region encoding a
 CC 170-amino acid polypeptide (see W81555), ORF2 encodes a putative
 CC rib-specific reductase (see W81554), ORF3 encodes a putative GTP
 CC cyclohydrolase (see W81552), ORF4 encodes the riboflavin synthase
 CC alpha subunit (see W81551), ORF5 encodes a putative rib-specific
 CC deaminase (see W81550) and ORF6 encodes a 38 kDa protein (see
 CC W81549) of unknown function. The invention provides vectors and
 CC recombinant bacteria for overproducing riboflavin, in which nucleic
 CC acids overproducing riboflavin biosynthetic proteins are introduced
 CC into the chromosome of the host organism at multiple sites or in
 CC multiple copies. The rib operon, e.g. comprising ORFs 2-5 and
 CC the beta-riboflavin synthase gene, is used to make such recombinant
 CC bacteria. The rib operon can be deregulated by replacing its
 CC control regions with sequences that allow constitutive or
 CC unregulated expression. The recombinant bacteria produce at least
 CC 10 g/l riboflavin.
 XX
 SQ

Sequence 5567 BP; 1749 A; 1102 C; 1298 G; 1418 T; 0 other;

Query Match 11.7%; Score 120.4; DB 20; Length 5567;

Best Local Similarity 51.0%; Pred. No. 4,9e-19;

Matches 317; Conservative 0; Mismatches 296; Indels 9; Gaps 1;

QY 1 ATGGATTATGCGATTCACATTCGCAATATGTCACAGTCACAAAGAGTTATCCACCC 60
 DB 1116 atgaagctggtccttagatcttcggaagcagggcagagacagacccaatccgcctc 1175
 QY 61 GTTGGCGCTGTTGTAAGTATGTAAGGTAGGATGTTGTTGCTGCACACTTGAGAAA 120
 DB 1176 gtccggtcgttgcgtlaaagcagcaaatgtcgcgaatggtgcgccattlaaataat 1235
 QY 121 GGTGACACATGCGGAGGTTCAAGCATTGATATGCGACACAAATATGCTGAAGTGGG 180
 DB 1236 ggtgaagcctcagcagaatctcgtccatccatgctcggagcaaatgcaagagtgcc 1295
 QY 181 ACGATTATATTACGTTAGTAGTATGATGATGTTGTTGTTCAACACACCCTGTGTTAAC 240
 DB 1296 gacatttaactgaactcgaaccgtgcagccacttaagcaaaaacacgcgcattgcaga 1355
 QY 241 AAAATTATGATTGTAGATAGCAAAAGTAGTATACGACAAAGAAAA-----CAAT 291
 DB 1356 ttgatatacaactcgtatctcaaaagagtgctgcgtatgagatatactaatccgctt 1415
 QY 292 TCGTTAGACACATGCTGATGACAGCTTACGGGCTTCACGGATTTAGGTTGAATCGTT 351
 DB 1416 gtggtcgtgaagagggatcagcatgaaagagcgtgcatgagtgaaagagcattc 1475
 QY 352 GATGATGACGCGGATCACACATATATACCAAGACTTTTAAAGCAAAAGCAAAAGCAACTG 411
 DB 1476 ctggcagaccagcgagagggcgaatgaaatctctcaactttagaggaagagccctt 1535
 QY 412 CCACAAATTACAGTGAAGATATCTGCAAGTTTATGATGTTAAACGCAATGATTAATGGA 471

DB 1536 ccgtacgtcacgtaaaagcggtgccagccttcagcgcaagataagctaccagcagcggc 1595
 QY 472 CAACTCAATGGATTACTTAACAAGAGGTAAACAGATGCTATTAAGTTAAGACATCGA 531
 DB 1596 gacagcaaatgtatcatcagctcagagcgtcgaagacagatgtctcgaatatacaggaaca 1655
 QY 532 CACGACGCACTGTTAACTGAAGACGTACAGTTGATATGATCCCAATATACTACA 591
 DB 1656 caccagaagcattttagctgcgagtcgcacagtgaaagcgcaacatccgacttaaccctgc 1715
 QY 592 CGTATTCAGATGCAAAAAC 613
 DB 1716 agactgcggaatgtacaacaac 1737

RESULT 13

ID 293752
 Z93752 standard; DNA: 5567 BP.

AC 293752;

DT 16-AUG-2000 (first entry)

XX Rib operon of *Bacillus subtilis*.

XX Riboflavin; rib operon; yeast; bacteria; transformation;

KW transcription element; gene expression; hair loss; skin inflammation;

KW growth disorder; vision; vision; sight; ds.

OS *Bacillus subtilis*.

FH Key Location/Qualifiers

FT CDS 1..5567

FT CDS 1084..1095

FT terminator

FT promoter

FT terminator

FT promoter

FT terminator

FT promoter

FT terminator

FT promoter

FT terminator

FT promoter

FT terminator

FT promoter

FT terminator

FT promoter

FT terminator

FT promoter

FT terminator

FT promoter

FT terminator

FT promoter

```
FT      /product= Beta riboflavin synthase
FT      4545..4574
FT      /tag= g
FT      /label= p3 promoter
FT      4653..4659
FT      RBS
FT      /tag= r
FT      4665..5039
FT      CDS
FT      /tag= s
FT      /label= ORF2
FT      5038..5090
FT      terminator
FT      /tag= t
FT      /label= Rho independent transcription terminator
FT      CDS
FT      complement (5053..5567)
FT      /tag= u
FT      /label= ORF1
FT      EPI001026-A2.
PN      17-MAY-2000.
XX      PD
XX      22-JUN-1990: 99EP-0124888.
XX      PF
XX      22-JUN-1989: 89US-0370378.
XX      PR
XX      22-JUN-1990: 90EP-0111916.
XX      PA
XX      (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PI      Perkins JB, Pero JG, Sloma A:
DR      WPI: 2000-320439/28.
DR      P-PSDB: Y83269, Y83270, Y83271.
XX      PT
XX      Novel vector, used for the large scale production of riboflavin,
XX      comprises a bacterial or yeast riboflavin biosynthetic protein under
XX      the control of a foreign transcription element
XX      Claim 1: Figure 3: 66pp: English.
XX      Vectors comprising a nucleic acid sequence of bacterial or
XX      yeast origin, coding for one or more riboflavin biosynthetic proteins,
XX      and one or more transcription elements not naturally associated with
XX      the nucleic acid sequence are new. The vector and bacteria transformed
XX      with it are used for the large scale production (over 10g/l) of
XX      riboflavin. The riboflavin produced can be used to treat disorders
XX      associated with riboflavin deficiency such as hair loss, skin
XX      inflammation, vision deterioration and growth failure.
XX      Sequence 5567 BP: 1751 A; 1097 C; 1296 G; 1423 T; 0 other:
SQ
Query Match 11.7%; Score 120.4; DB 21: Length 5567;
Best Local Similarity 51.0%; Pred. No. 4.9e-19;
Matches 317; Conservative 0; Mismatches 296; Indels 9; Gaps 1;
QY      1 ATGGATTATGCGATTGCAATATGTTACAGGTCAAGTCAACAGCGTGTAAACACC 60
DB      1116 atgaagctgagccttagatcttgcgaagcagggcgaagacagacccaatccgcctc 1175
QY      61 GTTGGCGTGTGTAGTAAAGAGTAGATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 120
DB      1176 gtgcggcgctgtgtcgttaagcagcaaatgtcgcgaatggcgccatttaaatat 1235
QY      121 GGTGACACGATCGGAGTTCACGACTTGATATGCGCACACAAATGCTGAAGTGGC 180
DB      1236 ggtgaagcctcatgcagaatltcatgcatccatattgctggaagcacaatgcagaggtgccc 1295
QY      181 AGATTATATACGTATGACGTATGACGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 240
DB      1296 gacattacgttaccatcgaaacgcgtgcacattacgsgaaacacacgcgcattgcagaa 1355
QY      241 AAAATTATGATTGTAAGTAGCAAAAGTAGATATACGCAACAAAGA-----CAAT 291
DB      1356 ttgattatcaactctgtatcaaaagagtgctgtgcgatgagagatcctaatccgctt 1415
```

```
QY      292 TCGTTAGACACATGTTGATGAGACGTTACGGCTCACGGTATTGAGTTGAATGCCGT 351
DB      1416 gtggtcgtggaagagggatcgcctgatgaaagaagctgtgcattgagtaaggaaagcctc 1475
QY      352 GATGATGAACGGGCATCACAAATTATACAGAGCTTTTAAAGCAAAAGCAAGCACTG 411
DB      1476 ctggcagaccagcgagagcgctgcatgaaanaattcttcactttatgagcagcgctt 1535
QY      412 CCACAAATTACAGTGAAGTATCTGCAAGTTTAGATGCTGAACCAAGCGAATATATGGA 471
DB      1536 ccgtacgtaccctaaagcgtgtccagccttgacgggaagatagctccagcagcggt 1595
QY      472 CAAAGTCATGATTACTTACAAAGAGGTTAAACAGATGCTATATAGTTAGACATCGA 531
DB      1596 gacagcaaatgtatcaccgtccagagctgcgaagacdgatgtccagcaataccggaataa 1655
QY      532 CACGACGAGTGTAACTGGAAGAGCTTACAGTTGAATTGATGATCCACAAATATCTACA 591
DB      1656 caccaaagcattttagtcgagtggtgcacagtgaaagccgacaatccgagcttaaccctgc 1715
QY      592 CGTATTCAGAGATGGAATAAAC 613
DB      1716 agactgcccgaatgttaacaaac 1737
RESULT 14
V52173/c
ID      V52173 standard; DNA: 14273 BP.
AC      V52173:
XX      23-OCT-1998 (first entry)
DE      Streptococcus pneumoniae genome fragment SEQ ID NO:40.
XX      Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
XX      computer readable medium; vaccine; pharmaceutical composition; ds.
XX      Streptococcus pneumoniae.
OS      WO9818931-A2.
XX      PN
XX      07-MAY-1998.
XX      PD
XX      30-OCT-1997: 97WO-US19588.
XX      PE
XX      31-OCT-1996: 96US-0029960.
XX      PR
XX      (HUMA-) HUMAN GENOME SCI INC.
XX      PA
XX      Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;
XX      Kunisch CA, Rosen CA:
XX      WPI: 1998-272225/24.
XX      DR
XX      Computer-readable medium with recorded Streptococcus pneumoniae
XX      polynucleotide sequences - useful in diagnostic kits and assays, and
XX      pharmaceutical compositions and vaccines for Streptococcus
XX      pneumoniae
XX      Claim 1: Page 389-398; 1409pp: English.
XX      PS
XX      The present invention describes a computer readable medium which has
XX      the nucleotide sequences SEQ ID NO:1 to 391 (V52134 to V52524) recorded
XX      on it, or a representative fragment or a sequence at least 95% identical
XX      to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in SEQ ID NO:1
XX      to 391 (V52134 to V52524) are genomic fragments from Streptococcus
XX      pneumoniae. The present invention also describes an isolated nucleic acid
XX      molecule encoding a homologue of any of the fragments of the S.pneumoniae
XX      genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced
XX      by a process comprising: (a) screening a genomic DNA library using as a
XX      probe a target sequence defined by any of the sequences in SEQ ID NO:1
```



```
OY 121 GGTGACAGCATGCGAGGATTCACAGCACTTGATGCGACACAAATGCTGAAGTGG 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 139 gggggcccaatgctgagagaataatgctcttaaaacgtgagaaatccctgctggagcg 198
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 181 ACGATTTATATTAGCTTACAGCCATGTAGTATTGTTGTTACACACACCCTGTGTTAAC 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 199 acgcttatgtaaacactgaaacccctgtgctacctcggaaacacccctcctgtagat 258
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 241 AAATTTGTGATGTAGATGCAAAAGTAGTATACCAACAAAGA-----CAATT 292
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 259 gctataatcgtatgtagtatacaagagtagtcaatggaagcctagactgtaatccatt 318
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 293 CGTTAGACACATGATGATGAGACGTTACGGGCTCACGGTATTGAGTTGATGCGTTG 352
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 319 gtatctgaaaaaggaagataatcttgagaaaaataatctcaagtactgtggaatt 378
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 353 ATGATGAACGGGCAATCAATTTATACCAAGACTTTTAAAGCA-AAAGCAAGCAACTG 411
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 379 ttaaaaaatgagtgcttaacttaataaaaaagtttagaaagtataatccagcatgta 438
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 412 CCACAAATTACAGTGAAGTATCTGCAAGTTTATGATGTAACACGCAATGATTAATGA 471
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 439 ccctatgtttttagaataatgcaatgcaatgtaaaaaatagccactaaacaat 498
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 472 CAAAGTCAATGATTTACTTAACAAGAGGTTTAAACAGATGCTTAAAGTTTACATGCA 531
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 499 caatccaaatggaatgactgaagaagaagaagaagcatgcatcagttcagcaactat 558
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 532 CACGACGCAAGTGTAACTGGAAGACGTACAGTTGAATTAGATGCCAATATATCTACA 591
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 559 gtagtgcaattatggtggagatcactacgttatacaagacgactccttgctgacatgt 618
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 592 CGTATTCAAGATGCAAAAACCTATATAAGTATATTGTCTAGTCTGGAAATATTCAT 651
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 619 agattggaagaaggaataacccatccgtatcatatcgtacacacattacgaacctcct 678
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 652 TTTA-ATCAGCAAAATTTATCAGATGATCAACACCAATTTGATATATCTGAAAAATCC 710
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 679 cttaaccttaaaatcgtaaaaaacgcaaatgataataaactacatgccaactcctc 738
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 711 AAATTTAACAAGCAATCAACACATATTGAATTTACTTGAAGTCTGTGATTTAAC 770
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 739 gaagacaaaataaaatgaagctatatacaaatcatgctgtgaatacttccataaag 798
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 771 AACAA 775
    |||||
Db 799 aaaaa 803
    |||||
```

Search completed: November 19, 2000, 05:07:08
Job time: 6894 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 02:32:43 ; Search time 111.68 Seconds
(without alignments)
1393.512 Million cell updates/sec

Title: US-08-978-456-1

Perfect score: 1029
Sequence: 1 ATGCATTATGCATTCACT.....TACTTACGAAGAAGTGA 1029

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 262060 seqs, 75620727 residues

Total number of hits satisfying chosen parameters: 524120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/2/ina/5C_COMB.seq: *
3: /cgn2_6/prodata/2/ina/5D_COMB.seq: *
4: /cgn2_6/prodata/2/ina/5D_COMB.seq: *
5: /cgn2_6/prodata/2/ina/6_COMB.seq: *
6: /cgn2_6/prodata/2/ina/PCRTUS_COMB.seq: *
7: /cgn2_6/prodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1029	100.0	1029	US-08-978-456-1	Sequence 1, Appl 1
2	1029	100.0	1029	US-08-978-456-1	Sequence 1, Appl 1
3	1029	100.0	1029	US-08-978-456-1	Sequence 1, Appl 1
4	352.2	34.2	1269	US-08-978-456-3	Sequence 3, Appl 1
5	112.6	10.9	3417	US-08-978-458-7	Sequence 7, Appl 1
6	112.6	10.9	3417	US-08-978-458-7	Sequence 7, Appl 1
7	76.2	7.4	1230	US-08-741-327E-14	Sequence 14, Appl 1
8	55.6	5.4	7218	US-08-232-463-14	Sequence 14, Appl 1
9	51.2	5.0	3095	5231168-1	Patent No. 5231168
10	47.8	4.6	9636	US-08-323-170B-1	Sequence 1, Appl 1
11	46.4	4.5	5589	US-08-465-795-1	Sequence 1, Appl 1
12	43.4	4.2	19124	US-08-487-826B-13	Sequence 13, Appl 1
13	41.6	4.0	4766	PCT-US93-07261-10	Sequence 10, Appl 1
14	40.4	3.9	665	US-08-883-795A-36	Sequence 36, Appl 1
15	39.2	3.8	4430	US-08-918-914-2	Sequence 2, Appl 1
16	38.8	3.8	1186	US-08-731-722-5	Sequence 5, Appl 1
17	38.8	3.8	3781	US-08-688-988-5	Sequence 5, Appl 1
18	38.6	3.8	19124	US-08-487-826B-13	Sequence 13, Appl 1
19	38.6	3.8	2817	PCT-US93-05944-1	Sequence 1, Appl 1
20	38.2	3.7	2530	US-08-450-351-1	Sequence 1, Appl 1
21	38.2	3.7	2530	US-08-450-351-1	Sequence 1, Appl 1
22	38.2	3.7	4248	US-08-678-614-1	Sequence 1, Appl 1
23	38	3.7	1939	US-07-715-751B-2	Sequence 2, Appl 1
24	37.8	3.7	10968	US-08-680-327-2	Sequence 2, Appl 1
25	37.8	3.7	2058	US-08-749-391-1	Sequence 1, Appl 1
26	37.2	3.6	1258	US-08-335-518-1	Sequence 1, Appl 1

27	37	3.6	2237	2	US-08-487-135B-1	Sequence 1, Appl 1
28	37	3.6	2237	3	US-08-915-972A-1	Sequence 1, Appl 1
29	37	3.6	2237	4	US-09-177-909-1	Sequence 1, Appl 1
30	36.6	3.6	446	5	US-09-097-541-1	Sequence 1, Appl 1
31	36.6	3.6	3926	4	US-08-731-722-1	Sequence 1, Appl 1
32	36.6	3.6	3926	4	US-08-731-722-1	Sequence 1, Appl 1
33	36.6	3.6	8920	3	US-08-446-855A-1	Sequence 1, Appl 1
34	36.6	3.6	56516	4	US-08-996-306-1	Sequence 1, Appl 1
35	36.4	3.5	3850	5	US-08-329-799-34	Sequence 34, Appl 1
36	36	3.5	3659	1	US-08-220-151-72	Sequence 72, Appl 1
37	36	3.5	3659	1	US-08-303-124-12	Sequence 12, Appl 1
38	36	3.5	3659	1	US-08-413-118-72	Sequence 72, Appl 1
39	36	3.5	3659	2	US-08-224-657-48	Sequence 48, Appl 1
40	36	3.5	3659	3	US-08-184-009-77	Sequence 77, Appl 1
41	36	3.5	3659	3	US-08-417-210A-67	Sequence 67, Appl 1
42	36	3.5	3659	3	US-08-480-697B-12	Sequence 12, Appl 1
43	36	3.5	3659	4	US-08-458-356-77	Sequence 77, Appl 1
44	36	3.5	3659	5	US-08-473-446-72	Sequence 72, Appl 1
45	36	3.5	3660	2	US-08-475-063-25	Sequence 25, Appl 1

ALIGNMENTS

RESULT 1
US-08-978-456-1
Sequence 1, Application US/08978456
Patent No. 6010881
GENERAL INFORMATION:
APPLICANT: Palmer, Leslie M.
TITLE OF INVENTION: No. 6010881el r1bG
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: Dechert Price & Rhoads
STREET: 4000 Bell Atlantic tower, 1717 Arch stre
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,456
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Todd O
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50444-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2252
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1029 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-978-456-1
Query Match 100.0%; Score 1029; DB 5; Length 1029;
Best Local Similarity 100.0%; Pred. No. 2e-246;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ATGCATTATGCATTCACTGCAATATGTCACAGGTCAACAGGTGTTATCCACCC 60

Db	1	ATGATTTATGGATTTCAACTTGCAAATATGTCACAAAGTCACAAAGGTTTATCTCACC	60
Qy	61	GTTGGCGCTGTTAGTTAAATGAAGTAGGATTGTTGTAATGGTCACATTCGAGAAA	120
Db	61	GTTGGCGCTGTTAGTTAAATGAAGTAGGATTGTTGTAATGGTCACACTTCGAGAAA	120
Qy	121	GGTACAAAGCATGGCGAGGTTCCAAAGCATCTGATATGGCAACAATAATGCTCAAGTCG	180
Db	121	GGTACCAAGCATGGCGAGGTTCAAGCATCTGATATGGCAACAATAATGCTCAAGTCG	180
Qy	181	ACGATTTATATTACGTAGAGACCATGATGATCTTTGGTTCAAACACACCTGCTGTAAAC	240
Db	181	ACGATTTATATTACGTAGAGACCATGATGATCTTTGGTTCAAACACACCTGCTGTAAAC	240
Qy	241	AAAAATTATGATTGTAAAGTAGCAAAAAGTAGTACGCCAACAAAGACAAATTCGTTAGAC	300
Db	241	AAAAATTATGATTGTAAAGTAGCAAAAAGTAGTACGCCAACAAAGACAAATTCGTTAGAC	300
Qy	301	ACACATGGTATGAGAGTTACGGGCTCAGGATATGAGTTGAATGGCTGATATGATGA	360
Db	301	ACACATGGTATGAGAGTTACGGGCTCAGGATATGAGTTGAATGGCTGATATGATGA	360
Qy	361	CGGGCATCACAAATTATACCAAGACTTTTAAAAAGCAAAAGCAAGCAATCGCCACAATT	420
Db	361	CGGGCATCACAAATTATATCCAAAGACTTTTAAAAAGCAAAAGCAAGCAATCGCCACAATT	420
Qy	421	ACAAGTAAAGTATTCGCAAGTTTAGATGCTAAACAACCGAATGATATGAGACAAAGTCAA	480
Db	421	ACAAGTAAAGTATTCGCAAGTTTAGATGCTAAACAACCGAATGATATGAGACAAAGTCAA	480
Qy	481	TGGATTTCTAACAAAGAGGTTAAACAAGATGCTATATAGTTAAGATGACAGACAGCA	540
Db	481	TGGATTTCTAACAAAGAGGTTAAACAAGATGCTATATAGTTAAGATGACAGACAGCA	540
Qy	541	GCTTAACTGGAAACGCTACAGTTGATATGATGATCCAAATATCTACAGTATTCGA	600
Db	541	GCTTAACTGGAAACGCTACAGTTGATATGATGATCCAAATATCTACAGTATTCGA	600
Qy	601	GATGAAAAAAACCTATPAAAAAGTATTTGTCTAAGCTGGGAAATATCTATTTAATCAG	660
Db	601	GATGAAAAAAACCTATPAAAAAGTATTTGTCTAAGCTGGGAAATATCTATTTAATCAG	660
Qy	661	CAATTTATCAAGTATGAAATCAACCCAAATTTGATATATACGAAAATCCAAATTTACA	720
Db	661	CAATTTATCAAGTATGAAATCAACCCAAATTTGATATATACGAAAATCCAAATTTACA	720
Qy	721	AGCATCAAAACACATATTGAATTTATTTACTTGAAGCTTTGGATTTAACAACAATTCCT	780
Db	721	AGCATCAAAACACATATTGAATTTATTTACTTGAAGCTTTGGATTTAACCAATTCCT	780
Qy	781	CACATTTATATTAAGAGAGAGGTTGGAATCTTGCTAGTGGAGGAGGTCCAAACCACTACT	840
Db	781	CACATTTATATTAAGAGAGAGGTTGGAATCTTGCTAGTGGAGGAGGTCCAAACCACTACT	840
Qy	841	TCGAATTTCTCCATCTATTTATATGATGATTTATCTCTATATTAATGCCCGCAATTAAT	900
Db	841	TCGAATTTCTCCATCTATTTATATGATGATTTATCTCTATATTAATGCCCGCAATTAAT	900
Qy	901	GCGCGATCTGGAAATTTATCAATTTTATCAACAACAAATGATGTGATGAGATACAGATCG	960
Db	901	GCGCGATCTGGAAATTTATCAATTTTATCAACAACAATGATGTGATGAGATACAGATCG	960
Qy	961	AACCAATTTGAAATGTTCAATTCGAGATTTAATAATCAAAATGTTAATTAATTAATTAACGA	1020
Db	961	AACCAATTTGAAATGTTCAATTCGAGATTTAATAATCAAAATGTTAATTAATTAATTAACGA	1020
Qy	1021	AAGAAGTGA 1029	
Db	1021	AAGAAGTGA 1029	

```

US-08-977-554-7
: Sequence 7, Application US/08977554
: Patent No. 5891672
: GENERAL INFORMATION:
: APPLICANT: Palmer, Leslie M.
: APPLICANT: Fedon, Jason C.
: APPLICANT: Warren, Richard L.
: APPLICANT: Traini, Christopher M.
: APPLICANT: Wang, Min
: APPLICANT: Jaworski, Deborah D.
: APPLICANT: Mooney, Jeffrey
: APPLICANT: Debowck, Christine
: APPLICANT: Zhong, Yixi
: APPLICANT: Black, Michael
: TITLE OF INVENTION: IIDA
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechert, Price & Rhoads
: STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
: CITY: Philadelphia
: STATE: PA
: COUNTRY: USA
: ZIP: 19103-2793
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSEO for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/977,554
: FILING DATE:
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US97/02318
: FILING DATE: 19-FEB-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Dickinson, O. Todd
: REGISTRATION NUMBER: 28,354
: REFERENCE/DOCKET NUMBER: P5044-07
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 215/994-2252
: TELEFAX: 215/994-2222
: TELEX:
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3336 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
:
: US-08-977-554-7
Query Match 100.0%; Score 1029; DB 3; Length 3336;
Best Local Similarity 100.0%; Pred. No. 2,9e-246; Indels 0; Gaps 0;
Matches 1029; Conservative 0; Mismatches 0;
QY 1 ATGATTATGCGATTGCAACTGCAAAATATGGTGACAGGTCAAAACAGGTGTTAATCCACC 60
DB 1 ATGATTATGCGATTGCAACTGCAAAATATGGTGACAGGTCAAAACAGGTGTTAATCCACC 60
QY 61 GTTGCGCGTGTGTAGTTAATGAAGGTGAGATGTGTGTTATTTGGTGACACTTGAGAAA 120
DB 61 GTTGCGCGTGTGTAGTTAATGAAGGTGAGATGTGTGTTATTTGGTGACACTTGAGAAA 120
QY 121 GGTGCAAGCATGGCGAGGTTCAGCACTGTGATATGGCACAACAAATGCTGAAGTGGC 180
DB 121 GGTGCAAGCATGGCGAGGTTCAGCACTGTGATATGGCACAACAAATGCTGAAGTGGC 180
QY 181 ACGATTATATACCTTGAAGCCATGTAGTATTTTGGTTCACACCCCTGTGTTAAC 240
DB 181 ACGATTATATACCTTGAAGCCATGTAGTATTTTGGTTCACACCCCTGTGTTAAC 240
QY 241 AAAATTATGATTGTAGATAGCAAAAGTAGTATTCGACCAAAAAGACATTCGTTAGAC 300

```



```
|||||
Db 241 AAAATATTGATGTGAAGTAGCAAAAGTAGTATACGCAACAAAGCAATTCGTTAGAC 300
Qy 301 ACACATGGTGATGAGACGCTTACGGGCTCACGGTATTGAGGTTAAATCGGTGATGAA 360
Db 301 ACACATGGTGATGAGACGCTTACGGGCTCACGGTATTGAGGTTAAATCGGTGATGAA 360
Qy 361 CGGGCATCAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCAAGTCCACAATTT 420
Db 361 CGGGCATCAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCAAGTCCACAATTT 420
Qy 421 ACAGTGAAGTATCTGCAAGTTAGTGTAAACAAGCGAATGATATGACAAAGTCAA 480
Db 421 ACAGTGAAGTATCTGCAAGTTAGTGTAAACAAGCGAATGATATGACAAAGTCAA 480
Qy 481 TGGATTACTAACAGAGGTTAAACAAGATGCTATTAAGTTAAGACATGCACGACGA 540
Db 481 TGGATTACTAACAGAGGTTAAACAAGATGCTATTAAGTTAAGACATGCACGACGA 540
Qy 541 GTGTTAACTGGAAGAGTACAGTTGAATTAGATGATCCAAATATCTACAGTATTCAA 600
Db 541 GTGTTAACTGGAAGAGTACAGTTGAATTAGATGATCCAAATATCTACAGTATTCAA 600
Qy 601 GATGGAAGAAACCCCTATTAAGATATGCTAAGTCTGGGAATATTCATTTTAACAG 660
Db 601 GATGGAAGAAACCCCTATTAAGATATGCTAAGTCTGGGAATATTCATTTTAACAG 660
Qy 661 CAATTTATCAAGATGAATCAACCAATTTGGATATATCTAATAATCCAAATTTACA 720
Db 661 CAATTTATCAAGATGAATCAACCAATTTGGATATATCTAATAATCCAAATTTACA 720
Qy 721 AGCAATCAACACATATTTGAATTTATTTACTGAGCTGTGATTTAACAAATTCCT 780
Db 721 AGCAATCAACACATATTTGAATTTATTTACTGAGCTGTGATTTAACAAATTCCT 780
Qy 781 CACAATTTATTAAGAGAGAGTGGAACTTCTAGTCAGGAGGAGTCCAACTACT 840
Db 781 CACAATTTATTAAGAGAGAGTGGAACTTCTAGTCAGGAGGAGTCCAACTACT 840
Qy 841 TCAGATTTCCATCTATTAATAGATGAATTTATCTATATATGCCCCGAATTAAT 900
Db 841 TCAGATTTCCATCTATTAATAGATGAATTTATCTATATATGCCCCGAATTAAT 900
Qy 901 GGGGATCTGGAATATCAATTTTATCAACAAATGATGATGATGATGATGATGCG 960
Db 901 GGGGATCTGGAATATCAATTTTATCAACAAATGATGATGATGATGATGATGCG 960
Qy 961 AACCAATTTGAATTTGATTCGAGTTTAAATCAAAATGTTAAATTAATTTACGA 1020
Db 961 AACCAATTTGAATTTGATTCGAGTTTAAATCAAAATGTTAAATTAATTTACGA 1020
Qy 1021 AAGAACTGA 1029
Db 1021 AAGAACTGA 1029

RESULT 3
US-08-978-456-7
; Sequence 7, Application US/08978456
; Patent No. 6010881
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; TITLE OF INVENTION: No. 6010881el r1bG
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```

```
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978, 456
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Todd O
REGISTRATION NUMBER: 28, 354
REFERENCE/DOCKET NUMBER: P50444-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2252
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3336 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-978-456-7

Query Match 100.0%; Score 1029; DB 5; Length 3336;
Best Local Similarity 100.0%; Pred. No. 2,9e+246;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGTATTGCGATTCACCTTCAATATGATGACAAAGTGAACAGAGTGTATCCACC 60
Db 1 ATGAGTATTGCGATTCACCTTCAATATGATGACAAAGTGAACAGAGTGTATCCACC 60
Qy 61 GTTGGCGTGTGTAGTAAATGAAGGTAGGATTTGTTGGTGCACACTTGAGAAA 120
Db 61 GTTGGCGTGTGTAGTAAATGAAGGTAGGATTTGTTGGTGCACACTTGAGAAA 120
Qy 121 GGTGAACACGCGGAGGTTCAAGCAGCTTGATATGGCAACAAATGCTAAGGTGG 180
Db 121 GGTGAACACGCGGAGGTTCAAGCAGCTTGATATGGCAACAAATGCTAAGGTGG 180
Qy 181 AGCATTTATTAAGTTAAGCATGTAGTCAATTTGGTTCAACACCACTGTGTAAAC 240
Db 181 AGCATTTATTAAGTTAAGCATGTAGTCAATTTGGTTCAACACCACTGTGTAAAC 240
Qy 241 AAAATTTGATTTGAATAGATAGCAAAAGTATAGCAAAAGCAAAAGCAATTCGTTAGAC 300
Db 241 AAAATTTGATTTGAATAGATAGCAAAAGTATAGCAAAAGCAAAAGCAATTCGTTAGAC 300
Qy 301 ACACATGGTGATGAGACGCTTACGGGCTCACGGTATTGAGGTTGAATGCGTTAGAA 360
Db 301 ACACATGGTGATGAGACGCTTACGGGCTCACGGTATTGAGGTTGAATGCGTTAGAA 360
Qy 361 CGGGCATCAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCAAGTCCACAATTT 420
Db 361 CGGGCATCAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCAAGTCCACAATTT 420
Qy 421 ACAGTGAAGTATCTGCAAGTTAGTGTAAACAAGCGAATGATATGACAAAGTCAA 480
Db 421 ACAGTGAAGTATCTGCAAGTTAGTGTAAACAAGCGAATGATATGACAAAGTCAA 480
Qy 481 TGGATTACTAACAGAGGTTAAACAAGATGCTATTAAGTTAAGACATGCACGACGA 540
Db 481 TGGATTACTAACAGAGGTTAAACAAGATGCTATTAAGTTAAGACATGCACGACGA 540
Qy 541 GTGTTAACTGGAAGAGTACAGTTGAATTAGATGATCCAAATATCTACAGTATTCAA 600
Db 541 GTGTTAACTGGAAGAGTACAGTTGAATTAGATGATCCAAATATCTACAGTATTCAA 600
Qy 601 GATGGAAGAAACCCCTATTAAGATATGCTAAGTCTGGGAATATTCATTTTAACAG 660
Db 601 GATGGAAGAAACCCCTATTAAGATATGCTAAGTCTGGGAATATTCATTTTAACAG 660
```

Db 601 GATGAAAAAACCCTATAAAGTAATATGCTCTAGCTGGGAATATTCATTTAATCAG 660
QY 661 CAAATTATCAAGATGATCAACACCAATTTGATATATGTAATATCAATTTAACA 720
Db 661 CAAATTATCAAGATGATCAACACCAATTTGATATATGTAATATCAATTTAACA 720
QY 721 AGCAATCAACACATATGTAATATGTAATATGTAATATGTAATATGTAATATGTA 780
Db 721 AGCAATCAACACATATGTAATATGTAATATGTAATATGTAATATGTAATATGTA 780
QY 781 CACATTTATATATAAGAGAGAGTGGAGCTTGTAGTGCAGGCGGTCCAACTACT 840
Db 781 CACATTTATATATAAGAGAGAGTGGAGCTTGTAGTGCAGGCGGTCCAACTACT 840
QY 841 TCAGATCTCCATCTATATATATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 841 TCAGATCTCCATCTATATATATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 901 GCGGATCTGGAATATATCAATTTATCAAAACAAATGATGATGATGATGATGATGAT 960
Db 901 GCGGATCTGGAATATATCAATTTATCAAAACAAATGATGATGATGATGATGATGAT 960
QY 961 AACCAATTTGAATTTGTCATTCGAGTTATTAATCAAAATGTTAAATTAATTA 1020
Db 961 AACCAATTTGAATTTGTCATTCGAGTTATTAATCAAAATGTTAAATTAATTA 1020
QY 1021 AAGAAGTGA 1029
Db 1021 AAGAAGTGA 1029

RESULT 4
: Sequence 3, Application US/08978456
: Patent No. 6010881
: GENERAL INFORMATION:
: APPLICANT: Palmer, Leslie M.
: TITLE OF INVENTION: No. 6010881el ribg
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechert Price & Rhoads
: STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
: CITY: Philadelphia
: STATE: PA
: COUNTRY: US
: ZIP: 19103
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/978,456
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Dickinson, Todd Q
: REGISTRATION NUMBER: 28,354
: REFERENCE/DOCKET NUMBER: P50444-9
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 215-994-2252
: TELEFAX: 215-994-2222
: TELEX:
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1269 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: US-08-978-456-3

Query Match 34.2%; Score 352.2; DB 5; Length 1269;
Best Local Similarity 88.4%; Pred. No. 1,1e-78;
Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;

QY 1 ATGATTAATGCGATTCACTTGCATTAATATGATGATGATGATGATGATGATGATGAT 60
Db 734 ATGATTAATGCGATTCACTTGCATTAATATGATGATGATGATGATGATGATGATGAT 793
QY 61 GTTGCGCTGTGTAGTATTAATGAAGTAGAGTGTGTGTATGTGTGTGTGTGTGTGT 120
Db 794 GTTGCGCTGTGTAGTATTAATGAAGTAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 853
QY 121 GGTGCAAGCATGCGGAGGTTCAGCACTGATGATGATGATGATGATGATGATGATGAT 180
Db 854 GGTGCAAGCATGCGGAGGTTCAGCACTGATGATGATGATGATGATGATGATGATGAT 913
QY 181 ACGATTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
Db 914 ACGATTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 973
QY 241 AAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 298
Db 974 AAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1033
QY 299 ACACACAT-GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 351
Db 1034 ACACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1093
QY 352 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 408
Db 1094 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1153
QY 409 CTGGCAC--AATTAACGTGAAGTATC-TGCAAGTTAGATGATGATGATGATGAT 462
Db 1154 CTGGCACAAATTAACGTGAAGTATC-TGCAAGTTAGATGATGATGATGATGAT 1213
QY 463 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 517
Db 1214 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1268

RESULT 5
: Sequence 7, Application US/08978458
: Patent No. 5932701
: GENERAL INFORMATION:
: APPLICANT: Palmer, Leslie M.
: APPLICANT: Fedon, Jason C.
: APPLICANT: Warren, Richard L.
: APPLICANT: Kosmatka, Anna L.
: APPLICANT: Shilling, Lisa K.
: APPLICANT: Stodola, Robert K.
: APPLICANT: Knowles, David J. C.
: APPLICANT: Black, Michael T.
: APPLICANT: Hodgson, John E.
: APPLICANT: Nicholas, Richard O.
: TITLE OF INVENTION: r1ba
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechert, Price & Rhoads
: STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
: CITY: Philadelphia
: STATE: PA
: COUNTRY: USA
: ZIP: 19103-2793
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/978,458
FILING DATE: 15-AUG-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/911,503
FILING DATE: 15-AUG-1997
APPLICATION NUMBER: PCT/US97/14436
FILING DATE: 15-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Q. Todd
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50533-09
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215/994-2252
TELEFAX: 215/994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3417 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-978-456-7

Query Match 10.9% Score 112.6: DB 4: Length 3417:
Best Local Similarity 49.2% Pred. No. 3.4e-19:
Matches 386: Conservative 0: Mismatches 389: Indels 10: Gaps 3:

QY 1 ATGATTTATGCGATTTCACCTTGCACAAATATGGTACAGGTCACAGGTTGTAATCCACC 60
DB 19 ATGAATATGACATAAATAAGGCGGCTGTTACGTCATCCCATTCATATG 78
QY 61 GTTGCGCGCTGTGTAGTAAATGAAGTACGATTTGTGTATGTCACACTTGAAGAAA 120
DB 79 GTTGCGCGCAATATGTAAGATATCATATTCGACAGGTTATCATGAGTTTTT 138
QY 121 GGTGACAGCATGGCGAGGTTCACAGCTTATATGGACAAACAAATGCGAAGGTGCG 180
DB 139 GGTGCGCCCATGCTGACAGAAATGCTTTAAAACTGTACAAAATCCCTGTGAGCG 198
QY 181 ACGATTTATATACGTAGAGCCATGATCTATTTGTTTCACACGACCCCTGTGTAC 240
DB 199 ACGTTTATGTAACACTGAACCTGTGTCATCTCGGAAACACCCCTGTATGAT 258
QY 241 AAAATTATGTTTGAATGCAAAAGTACTATACGCAACAAAGA-----CAATT 292
DB 259 GCTATATCGATAGTGTATACAGAGTAGTCAATGGAAACCTAGACTGTATCTATT 318
QY 293 CGTTAGACACATGATGAGAGGTTACGGGCTCAOGGATATGAGGTTGATGCGTTG 352
DB 319 GTATCTGAAAAAGAGTAAGATCTTGAGAAAAATATCTTCAAGTTACTGTGAATT 378
QY 353 ATGATGAACGGGATCACAATTTATACCAAGACTTTTTTAAAGCA-AAAGCAACCAATG 411
DB 379 TTGAAAAATGAGTCTTACTAATAAAAAAGTTTGAAGATATTAACCCACAGTA 438
QY 412 CCACAAATTACAGTGAAGTCTGCAAGTTTACATGTAACAGGAAATGATATGGA 471
DB 439 CCCATATGTTTTTGAATATGCAATGTCAATGATGTAAGCAACCAAT 498
QY 472 CAAAGTCAATGATTAACAAGAGTTAAACAAGATGCTTAAGTAAAGCATGCA 531
DB 499 CAATCCAAATGATTAACAAGAGGAAAGCAAGATGCAATGCAATGCACTAT 558
QY 532 CACGACGAGTGTAACTGTAAGAGTGAAGTGAATTAATGATGATCCATATATCTACA 591
DB 559 GTTAGTCATATGATGAGGAGTCAATATCTTATCAAGACGATCCTTGTGTCATG 618
QY 592 CGTATTCAGATGAAAAACCTATATAAAGTAAATTTGTTAAGTGTGGAAATATTCAT 651
DB 619 AGATTGAGAGGAGAAAAATCCTATCCGTATCATATGCGATACACATTTACGAATCCT 678

QY 652 TTAA-ATCAGCAATTTATCAAGATGAATCAACACCAATTTGGATATATATCAAAATCC 710
DB 679 CTTAACCTTAATAATCGTAAAAACAGCAAAATGATATTAACCTTACATTCGCACTTCTCT 738
QY 711 AAATTTACAGCAATCAACATCAATATTTGAATTTACTTGAAGTCTTGTGATTTAC 770
DB 739 GAAGCAAAAAATTAATGAAGCTATATCAAAATCATGAGTGTGAATATCTTCCATTAAG 798
QY 771 AACA 775
DB 799 AAAA 803

RESULT 6
US-08-978-454-7
Sequence 7, Application US/08978454
Patent No. 6017728
GENERAL INFORMATION:
APPLICANT: Palmer, Leslie M.
APPLICANT: Fedon, Jason C.
APPLICANT: Warren, Richard L.
APPLICANT: Kosmatka, Anna L.
APPLICANT: Shilling, Lisa K.
APPLICANT: Stodola, Robert K.
APPLICANT: Knowles, David J. C.
APPLICANT: Black, Michael T.
APPLICANT: Hodgson, John E.
APPLICANT: Nicholas, Richard O.
TITLE OF INVENTION: r1bH
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert, Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2793
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,454
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/911,503
FILING DATE: 15-AUG-1997
APPLICATION NUMBER: PCT/US97/14436
FILING DATE: 15-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Q. Todd
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50533-11
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215/994-2252
TELEFAX: 215/994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3417 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-978-454-7

Query Match 10.9% Score 112.6: DB 5: Length 3417:
Best Local Similarity 49.2% Pred. No. 3.4e-19:
Matches 386: Conservative 0: Mismatches 389: Indels 10: Gaps 3:
QY 1 ATGATTTATGCGATTTCACCTTGCACAAATATGGTACAGGTCACAGGTTGTAATCCACC 60

```

19 ATGAATATAGCAATAAACTGCACAAAAGGGGCTGTACGTCAATCCCAATCCTATG 78
61 GTTGCGCGCTGTGTATGATGAGGTAGGATGTTGTGATGTGCACACTTGAAGAAA 120
79 GTTGCGCGCAATTATGTTAAAGTAATACATTATGCGCAAGGTTATATATGATTTT 138
121 GGTGACACAGCATGCGAGGTTCAAGCACTTGATGCGCAACAAATCTGAGGTGCG 180
139 GGTGCGCCACATGCTGAGAGAAATGCTTTAAACTGTAGAAAATCCCTGCGAGGG 198
181 ACGATTATATTAATGCTTGAAGCCAGTGTATCATTTTGGTTCAACACCCTGTGTTAC 240
199 ACGCTTATGTAACTTAACCTTACCTGTTCTCACTTGGGAAAAACCTCCCTGATAGAT 258
241 AAATATATGATTTGTAGATAGCAAAAGTATAGCAACAAAAG-----CAAT 292
259 GCTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 318
293 CGTTAGACACACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 352
319 GTATCTGGAAGAGGATTAAGATCTGAGGAATAATCTTCAATCTGTTGGAAT 378
353 ATGATGACGGGCATCACAAATATACCAAGACTTTTAAAGCA-AAAGCAAGCACTG 411
379 TTAGAAAATGAGTGTCTTAATTAATAAAAGTTTGAAGAGTATTTACCGACGATGA 438
412 CCACAAATACAGTGAAGTATCTGCAATTTAGATTTAGTAAACAGCAATGATATGA 471
439 CCTATGTTTATATGAAATATGCAATGCAATGATGATGATGATGATGATGATGAT 498
472 CAAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 531
499 CAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 558
532 CACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 591
559 GTTAGTGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 618
592 CATTATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 651
619 AATATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 678
652 TTAT-ATCAGCAATTTATCAAGATGATGATGATGATGATGATGATGATGATGATG 710
679 CTATCTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 738
711 AAATTTACAGCAATCAACACATATGAAATTTATTTACTGAGTCTTGTGATTTAAC 770
739 GAAGACAAAATTAATGAAGTATATCAAAATCATGCTGTGAATACTTCCATTAAG 798
771 AACAA 775
799 AAAA 803

```

```

US-08-741-327E-14
; Sequence 14, Application US/08/41327E
; Patent No. 5925354
; GENERAL INFORMATION:
; APPLICANT: Fuller, et al.
; TITLE OF INVENTION: Riboflavin mutants as vaccines
; TITLE OF INVENTION: against Actinobacillus pleuropneumoniae
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: G. Kenneth Smith
; STREET: 300 S. Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/741.327E
; FILING DATE: October 28, 1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: G. Kenneth Smith
; REGISTRATION NUMBER: 43,135
; REFERENCE/DOCKET INFORMATION:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1230 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-741-327E-14

```

```

Query Match 7.4%, Score 76.2; DB 4; Length 1230;
Best Local Similarity 47.6%; Pred. No. 2,6e-10;
Matches 293; Conservative 0; Mismatches 313; Indels 9; Gaps 2;

```

```

9 TGGCATTCACCTTGCAAAATATGATGATGATGATGATGATGATGATGATGATGATG 68
165 TGGCATTCACCTTGCAAAATATGATGATGATGATGATGATGATGATGATGATGATG 224
69 TGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 128
225 TGTATTTGTCAAAACGTTGAATGCTTGGCGAAGTTACATGAAAGATTTGGTGATG 284
129 GCATCGGAGGTTCAAGCATTTGATGATGATGATGATGATGATGATGATGATGATG 188
285 GCATCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 344
189 TATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 248
345 TGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 404
249 TGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 305
405 TGAAGAGGATTAATAAAGTATTTATGCTTGCAGGATCCGAATCCTTAGTAGCAGG 464
306 TGGTG-----ATGAGACGTTACGGGCTACGCTATGATGATGATGATGATGATGATG 359
465 GCGGGAGCAAAATCAGCTACGCCAAGCCGCGTGAAGTGGGAGGATTTACTCAAGA 524
360 ACGGCATCACATATATCAAGACTTTTAAAGCAAAAGCAAGCAAGCAAGCAAGCA 419
525 AGAATGTATGCTTAAACCCGATTTTTCACATATTTCAACTAATCAAGCTCGTATGT 584
420 TACAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 479
585 GCTATGAAATATGCTACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 644
480 ATGATTAATCAAAAGAGGTTAAACAAAGATGCTATATAGTAAAGATGATGATGATG 539
645 ATGATTTACGGTGAATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 704
540 AGTGTAACTGGAAGACGTACAGTTGATGATGATGATGATGATGATGATGATGATG 599
705 GATTTATGCTGCTGTAGATAGGATCTTGCATTAACCCGATGTTAAATACCGAATGCC 764
600 AGATGAAAAAACCC 614
765 GAATCGAACAACC 779

```

RESULT 8
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 5.4%; Score 55.6; DB 1; Length 7218;
Best Local Similarity 7.2%; Pred. No. 5.8e-05;
Matches 31; Conservative 220; Mismatches 179; Indels 0; Gaps 0;

229 CCCTGTGTAACAAATTTGATGTTAGATGCAAAAGTAGTATACGCAACAAAGAC 288
1487 CATGTAATTAACCTATCTATGACAGTAGTTAAAGAGATGAGAAATTTGTACRRRRRR 1428
289 AATTCCTGTGACACACATGTGATGAGCTTACGGCTCAGCGTATGAGTTGAATG 348
1427 RRR 1368
349 GTTGATGATGACGGGCATACATTAATACCAAGACTTTTAAAGCAAAAGCAACAA 408
1367 RRR 1308
409 CGGCCCAATTAATGATGAGATGATGCAAGTTAGATGTTAAGCAAGCATGATAT 468
1307 RRR 1248
469 GCACAAGTCAATGATGATTAACAAGAGTTAAACAAGATGCTATATAGTTAAGACAT 528
1247 RRR 1188

529 CGACACGACGACGTGTTAAGTGAAGACGTACGTGTAATGATGATCCACATATACT 588
1187 RRR 1128
589 ACACGTATTCAGATGAGAAAAACCTATAAGTAAATTTGCTTAAGTCTGGCATATT 648
1127 RRR 1068
649 CATTTTAATC 658
1067 RATCGCAAGC 1058

RESULT 9
5231168-1
Patent No. 5231168
APPLICANT: DIEGIEL, MORTEN; BORRE, MARTIN; JEPSEN, SOREN;
VUUST, JENS; RIENECK, KLAUS; WIND, ANNETTE; JAKOBSEN, PALLE H.
TITLE OF INVENTION: MALARIA ANTIGEN
NUMBER OF SEQUENCES: 19
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/409,658
FILING DATE: 18-SEP-1989
SEQ ID NO: 1
LENGTH: 3095
5231168-1

Query Match 5.0%; Score 51.2; DB 7; Length 3095;
Best Local Similarity 41.6%; Pred. No. 0.00055;
Matches 323; Conservative 0; Mismatches 453; Indels 0; Gaps 0;

250 GATTGATGATGCAAAAGTATATACGCAACAAAGCAATTCCTTAGACACATGCT 309
1408 gaagataaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 1467
310 GATGAGACGTTACGGCTCAGCTGTTGATGATGATGATGATGATGATGATGATGAT 369
1468 gataaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 1527
370 CAATTATACCAAGCTTTTAAAGCAAAAGCAACCACTGCCAATTAATCACTGAAA 429
1528 aaaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 1587
430 GTATCTGCAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 489
1588 aatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 1647
490 AACAAAGAGCTTAAACAAAGATGCTATATAGATGATGATGATGATGATGATGATGAT 549
1648 aatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 1707
550 GGAAGACGTACAGTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 609
1708 gaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 1767
610 AACCTATATAAGTATATGCTTAAGTCTGGCAATATTCATTTTAATCAGCAAAATTTAT 669
1768 aagatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 1827
670 CAAGATGATCAACACCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 729
1828 gaagatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 1887
730 ACACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
1888 aagatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 1947
790 TATTAAGAGGAGTTGGAATTTGCTAGTGCAGGACGATGATGATGATGATGATGATGATGAT 849
1948 aatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 2007

QY 850 TCACCTTTTATGATGATTTTCTCTATTGGCCGAATTAATGGCGATC 909
 Db 2008 gtccaattaaatgtcacaagaataataaccaataaagaacccagtaatgat 2067
 QY 910 GGAATTTTCATTTTATCAACAATGATGATTGATCCAGATGCGAACCAATT 969
 Db 2068 ggtccaaacatgtgaacaataatcacagaagatgataatgatgaagaagatgat 2127
 QY 970 GAAATTTTCATTCGCGATTATTAATCAAAATGTTAAATTAATTAATTCGAAGA 1025
 Db 2128 gataatgatttgaagaattatcaagaagaagatgtgataaagaattcatcaatba 2183

RESULT 10
US-08-323-170B-1
; Sequence 1, Application US/08323170B

TITLE OF INVENTION: Cloning and Expression of Plasmidium
 TITLE OF INVENTION: falciiparum Transmission-Blocking Target Antigen, Pf5230
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, 8th Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/323,170B
 FILING DATE: 13-OCT-1994
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/010,409
 FILING DATE: 29-JAN-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Quine, Jonathan A.
 REGISTRATION NUMBER: P-41,261
 REFERENCE/DOCKET NUMBER: 015280-113100US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9636 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 149..9556
 US-08-323-170B-1

Query Match	4.6%	Score 47.8	DB 2	Length 9636
Best Local Similarity	49.0%	Pred. No. 0.0054		
Matches 127	Conservative	0	Mismatches 132	Indels 0
				Gaps 0

0y	460	ATATATATGACCAAAATCATGATGATCTACTACACAAAAGAGTTAAACCAAGATGCTCATAG	519
Db	6941	AGTATATTGAAAGTACTGAAACATATTATTTATTTATTCATTAAGATATAAAAATATATAT	7000
0y	520	TTAAGACATGCACACGCGATGTTAACTGGAAGACGTCACAGTTGAATTAGATATGCCA	579
Db	7001	TTAAACCTTAAGATGTTATTGAATTAATGATGAAGAAATTTTAGAGAAATTAACAAAATATA	7060

QY	580	C A R T A T C T A C G A T A T C A A G T G S A M A A A A C C C T A T A A A G T A T A T Y T G C T A A G C T	639
Db	7061	C A C A A T A T T A C A G S A T A T C A A A T A T T A C A G A T G T T A C A T T C A A A A A T T T A A T T A	7120
QY	640	G G S A A T T A T A T T T A T A T C A G C A A A T T P C A A G A T G A T C A A C C A A T T T G A T A T A T	699
Db	7121	G G T A T A T T A C C A T T A A A T T T T A A A A A T C A T T A T T C A C A G A T A T G C T A A A G T A C C A G A T	7180
QY	700	A C T G A A A T C C A A A T T T A A	718
Db	7181	A C C T T T A A T T C T A T T A T T A	7199

QY	700	ACTGAAATCCAAATTTAA	718
Db	7181	ACCTTAATCTATTATTA	7199

RESULT 11
US-08-465-795-1

APPLICANT: Yonetani, Yoshiyuki
 APPLICANT: Teshiba, Sadao
 TITLE OF INVENTION: A Process for Producing Riboflavin
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
 ADDRESSEE: Dunner
 STREET: 1300 I Street, N.W., Suite 700
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20005-3315
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/465,795
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/161,394
 FILING DATE: 06-DEC-1993
 APPLICATION NUMBER: JP 326578/1992
 FILING DATE: 07-DEC-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Gaybrick, Robert J.
 REGISTRATION NUMBER: 27,890
 REFERENCE/DOCKET NUMBER: 04853.0008-00000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-408-4000
 TELEFAX: 202-408-4400
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5589 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Corynebacterium ammoniagenes
 US-08-465-795-1

[illegible]

NAME: Blasdale, John H. C.
REGISTRATION NUMBER: 31,895
REFERENCE/DOCKET NUMBER: DX0288K
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-822-7398
TELEFAX: 201-822-7039
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 4766 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
STRAIN: Malayan Camp
IMMEDIATE SOURCE:
CLONE: p2bl:p12-1
FEATURE:
NAME/KEY: CDS
LOCATION: 3..4766
PCT-US93-07261-10

Query Match 4.0%; Score 41.6; DB 6; Length 4766;
Best Local Similarity 46.5%; Pred. No. 0.15;
Matches 134; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

OY 442 TTAGATGTAACAAGCGCATGATATGACAAAGTCATGCTACTACCAAGAGTT 501
DB 189 TTTAAAGAAAAAAGAAATTAAGAAACATTTAAAAAAGAGTTGAAATGAGAGAA 248
OY 502 AAACAGATGCTATAGTACGACACAGCGAGCTTTAACTGGAAGAGTACA 561
DB 249 AAAGAGAAAAAATTAATTAAGATATATGATGAGCATTTAAGATATAAGAAAT 308
OY 562 GTTGATTTAGATGATCCACATATATACGCTATTCAAGATGAGAAAAACCTATAAA 621
DB 309 GATTAATGATGATAAAAAATAGTCCCAAAAAACCAAGATCCGTGAAAAAGATTGAAA 368
OY 622 GTAATATGCTAAGTCTGGGAATATTCATTTTATCAGCAATTTTCAAGATGATCA 681
DB 369 GAAATGGAATGAGAGAGAAATTTATTAACAAACATTTAAGATTTATGAGAGAGA 428
OY 682 ACACCAATTTGATATATCTGAAAAATCCAAATTTAAGACGATCAA 729
DB 429 AAGGAAAAAAGAAATTTGATCTACTAGACACTTTAAGAGAGATAAA 476

RESULT 14
US-08-883-795A-36/c
Sequence 36, Application US/08883795A
Patent No. 5985607
GENERAL INFORMATION:
APPLICANT: Delcive, Genevieve
APPLICANT: Awang, Gregor
TITLE OF INVENTION: Recombinant DNA Molecules and Expression
TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,795A

FILING DATE: 27-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Micheline
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 7841-062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 665 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: Rh 32
US-08-883-795A-36

Query Match 3.9%; Score 40.4; DB 4; Length 665;
Best Local Similarity 44.6%; Pred. No. 0.16;
Matches 207; Conservative 0; Mismatches 251; Indels 6; Gaps 1;

OY 494 AAGAGCTTAACAAGATGCTATAGTAAAGACATCGACACGACGAGCGTGTACTGGA 553
DB 516 AAATATGTAATTTAATTAATCTTTAATTAATTAATTAATTAATTAATTAATTA 457
OY 554 GACGTACAGTGAATTTAGATGATGATGATGATGATGATGATGATGATGATGATG 613
DB 456 TATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 397
OY 614 CTATTAAGTATATTTGCTAAGTCTGGGAATATTCATTTTATCAGCAATTTATCA 673
DB 396 ATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 337
OY 674 ATGATACACACCAATTTGGATATATCTGAAATCCAAATTTTCAAGCAATCAACAC 733
DB 336 ATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 277
OY 734 ATAT-----TGAATTAATTTACTTGAAGTCTGATTTAACAACATTTCTTCAATT 787
DB 276 ATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 217
OY 788 TATATTAAGAGAGCTTGGAACTTGTCTAGTCGAGCAGGTCACACCTACTTCAGAA 847
DB 216 ATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 157
OY 848 TCTGCATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 907
DB 156 ATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 97
OY 908 CTGGAATTTATCAATTTTATCAACAAGATGATGATGATGATGATGATGATGATGAT 951
DB 96 ATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 53

RESULT 15
US-08-918-914-2/c
Sequence 2, Application US/08918914
Patent No. 5876963
GENERAL INFORMATION:
APPLICANT: Mitchell, Peter
APPLICANT: Hutchinson, Nancy
APPLICANT: Lawton, Michael
APPLICANT: Magna, Holly
APPLICANT: Yocum, Sue
APPLICANT: Murry, Lynn E.
TITLE OF INVENTION: HUMAN NUCLEOTIDE PYROPHOSPHORYLASE
NUMBER OF SEQUENCES: 4


```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/918,914
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0369
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4430 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: 777
CLONE: 422069
US-08-918-914-2

```

```

Query Match          3.8%; Score 39.2; DB 3; Length 4430;
Best Local Similarity 49.5%; Pred. No. 0.57;
Matches 101; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 670 CAAGATGATCAACACCAATTTGGATATATATACGAATCCAAATTTAACAAGCAATCAA 729
DB 4242 CTAATATTAAAGCTTAACCTTGATGACACTAGAAATTCGAATTTAGCAAGAAATAA 4183
QY 730 ACACATATTGAATTAATTTACTGAGCTCTTGATTTAACAACAATCTTCACAATTTA 789
DB 4182 ATAAGTTATTTCATTTCAAGTAAGTAATGTGTTTATTCTTTATTTCACCAACATT 4123
QY 790 TATATAAGAGAGTGGAACTTGCTAGTCGAGGCAAGTCCAACTACTTCAGAAATTC 849
DB 4122 TATATGTATGAACTAGCAAGGCTTGCTATGTCTGTGAGCAAGGCAAGAAATAT 4063
QY 850 TCCATCTATTATAGTAATTT 873
DB 4062 GCTTATATTGTTATTGATCTT 4039

```

Search completed: November 19, 2000, 05:00:11
Job time: 8848 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 00:12:31 ; Search time 1232.02 seconds
(without alignments)
5163.965 Million cell updates/sec

Title: US-08-978-456-1
1029
Sequence: 1 ATGATATATCGATCACT.....TAACCTTAGAAGAGATGA 1029

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7189864 seqs, 3091403243 residues

Total number of hits satisfying chosen parameters: 14379728

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: em_estba:*
42: em_estfun:*
43: em_esthum1:*

44: em_esthum2:*
45: em_esthum3:*
46: em_esthum4:*
47: em_esthum5:*
48: em_esthum6:*
49: em_esthum7:*
50: em_esthum8:*
51: em_esthum9:*
52: em_esthum10:*
53: em_esthum11:*
54: em_esthum12:*
55: em_esthum13:*
56: em_esthum14:*
57: em_esthum15:*
58: em_esthum16:*
59: em_esthum17:*
60: em_esthum18:*
61: em_esthum19:*
62: em_esthum20:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estom:*
68: em_estov1:*
69: em_estov2:*
70: em_estp11:*
71: em_estp12:*
72: em_estp13:*
73: em_estp14:*
74: em_estp15:*
75: em_estro1:*
76: em_estro2:*
77: em_estro3:*
78: em_estro4:*
79: em_estro5:*
80: em_estro6:*
81: em_estro7:*
82: em_estro8:*
83: em_estro9:*
84: em_estro10:*
85: em_estro11:*
86: em_estro12:*
87: em_estro13:*
88: gb_gss1:*
89: gb_gss2:*
90: gb_gss3:*
91: gb_gss4:*
92: em_gss1:*
93: em_gss2:*
94: em_gss3:*
95: em_gss4:*
96: gb_gss5:*
97: gb_gss6:*
98: gb_gss7:*
99: gb_gss8:*
100: gb_gss9:*
101: em_gss5:*
102: em_gss6:*
103: em_gss7:*
104: em_gss8:*
105: em_gss9:*
106: em_gss10:*
107: em_gss11:*
108: gb_gss10:*
109: gb_gss11:*
110: em_gss12:*
111: gb_gss12:*
112: gb_gss13:*
113: gb_gss14:*
114: gb_gss15:*
115: gb_gss16:*
116: gb_gss17:*

OY	990	AACAAATGATGTGCATGATGACATCCAGCATCGAACAATTGAATGGTCATTCGAGTT	989
Db	1002	AAAAAAAAATTMAAA---AAWMTTAAAMAATCAAAAAAATTTCTTTTCCAAAAM	1058
OY	990	ATTAATCAAAATGTTAAATTAATTAACCTTACCAGAAGAA	1025
Db	1059	WHAAAAAHMAYCMWTYTTCMATCTTTTMAAAA	1094
RESULT	2		
AV426745			
LOCUS	AV426745	420 bp	mRNA EST 23-MAY-2000
DEFINITION	AV426745	Lotus japonicus young plants (two-week old) Lotus japonicus cDNA clone MIM071b02_r_5', mRNA sequence.	
ACCESSION	AV426745		
VERSION	AV426745.1	GI:7785990	
KEYWORDS	EST.		
SOURCE	Lotus japonicus.		
ORGANISM	Lotus japonicus		
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta		
AUTHORS	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;		
TITLE	Fabales; Fabaceae; Papilionoideae; Lotus.		
JOURNAL	I (bases 1 to 420)		
MEDLINE	Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.		
COMMENT	Generation of 7137 non-redundant expressed sequence tags from a legume, Lotus japonicus		
FEATURES	Legume, Lotus japonicus		
Source	DNA Res. 7 (2), 127-130 (2000)		
Location/Qualifiers	Contact: Yasukazu Nakamura The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: ynakun@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/ Location/Qualifiers		
BASE COUNT	98 a 72 c 132 g 118 t		
ORIGIN	1..420		
	/organism="Lotus japonicus"		
	/db_xref="taxon:34305"		
	/clone="MM071b02_r"		
	/clone_lib="Lotus japonicus young plants (two-week old)"		
	/dev_stage="young plants (two-week old)"		
	/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:		
	XhoI; isolate=miyakoJima MG-20"		
	XhoI; isolate=miyakoJima MG-20"		
Query Match	6.6%, Score 68.2; DB 18; Length 420;		
Best Local Similarity	55.8%; Pred. No. 6.8e-06;		
Matches 130; Conservative	0; Mismatches 103; Indels 0; Gaps 0;		
OY	40	CAAACAGGCTTAATCCACCCGCTTGCGCGCTGTAGTAAAGAAAGTAGATTGTTGTT	99
Db	1	CAGACCAACCCCAATCCATGTGTGGGCTGTATATGTGAAGGATGGAAAAATGTTGTT	60
OY	100	ATTGCTGCACACTTGAGAAAAGTGACAAAGCATGCGGAGGTTCAAGCATTGATATGCCA	159
Db	61	CAAGGCTTACACCTTAAGCAGGACGCTCATGCTGAGAGCTTTTGCTTAAGAGATCGG	120
OY	160	CAACAAATGCTGAAGGTCGACGATTTATATAGCTTAGACCATGTAGTATTGTTGTT	219
Db	121	GCGGATTTGGCAGAGGTCGACGCTTATGTGAGTTTGAGACCTTGATATCATTTTGGT	180
OY	220	TCAAACACACCCCTGTGTTAACAAAATTAATGATTGATGAGCAAAAGTAGT	272
Db	181	AGGACTGCACCTTGTTCGAACTTAATTCAAGCCAAAGTGAAGAGGTGTT	233
RESULT	3		
LOCUS	CNS00396/c	1101 bp	DNA GSS 03-JUN-1999
DEFINITION	CNS00396/c		
	BACR08K10 of Rpci-98 library from Drosophila melanogaster (fruit		

Accession	Version	KeyWords	Source	Organism	Reference	Authors	Title	Journal
AL063921	1	GI:4941778	GSS	<i>Drosophila melanogaster</i>	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
COMMENT	<p>fly), genomic survey sequence.</p> <p>Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)</p> <p>Web : www.genoscope.cns.fr</p> <p>Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oseegawa and Aaron Mammoss in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by Paolo, ECORI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw^{sp}, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.</p> <p>Location/Qualifiers</p> <p>1. 1101</p> <p>/organism="Drosophila melanogaster"</p> <p>/db_xref="taxon:7227"</p> <p>/clone_id="RPCI-98"</p> <p>/clone="BACR08K10"</p> <p>/note="end : TET3"</p>							
FEATURES	<p>Source</p> <p>1. 1101</p> <p>/organism="Drosophila melanogaster"</p> <p>/db_xref="taxon:7227"</p> <p>/clone_id="RPCI-98"</p> <p>/clone="BACR08K10"</p> <p>/note="end : TET3"</p>							
BASE COUNT	<p>201 a 64 c 131 g 202 t 503 others</p>							
ORIGIN	<p>1101</p>							
Query Match	<p>6.1%; Score 63.2; DB 121; Length 1101;</p> <p>Best Local Similarity 17.4%; Pred. No. 0.0011;</p> <p>Matches 110; Conservative 276; Mismatch 244; Indels 4; Gaps 2.</p>							
Db	<p>189 TATTACGTTAGAGCCATGTAGTCATTTTGTTCAACACACCCTGTGTTAAACAATTAT 248</p> <p>1100 KARRGGDTWDRPTKDDMDWTWTKWKDRADRRMAGDADRMAWDGAGTWTATATWW 1041</p>							
Y	<p>249 TGAATGTAAAGTATGCAAAAGTAGTATACGCAACAACAAGCATTTGTTAGACACATATG 308</p> <p>1040 WWWWATATWTDWDMKWWMTAAKDTDTAATWMTARABWAGDRGAKGRDRDRAADADGCA 981</p>							
Y	<p>309 TGATGAGACGTTAACGGCGTACCGATGAGAGTGAATG--CGTGAAGATGAACGGCA 366</p> <p>980 GRDGGRRRKDKDKDDDDKGGKKKKAAMAAWATKMDWDDWDMKWDGAKDKRA 921</p>							
Y	<p>367 TCACATATATACCAAGATTTTTTAAACCAAAAGCAAGCAACTGCCAAATTACAGT 426</p> <p>920 DDDDSAGKDDDKGKADDDDTGKDDDKDKMDMDMKANGTWDATWMAATDMMW 861</p>							
Y	<p>427 AAAGTATGCAAGTTTATGATGTTAACAAGCAATGATTAATGCAAAATGATGATT 486</p> <p>860 WADADMTWTDAAADDMWADDKWDAAWAKWMDMAWAGARTADRDWDGAKGRKARRD 801</p>							
Y	<p>487 ACTAACAAGAGGTTAAACAAGATGTCATATAAGTTAAGACATCGACAGACGATGTTA 546</p> <p>800 KRRADKRDADDDRDAAATWTTTTRTTTTRTDDWKKWKTDTWRMAADTRPWRDDDDSDR 741</p>							
Y	<p>547 ACT--GGAAGACGTCAGTGAATATAGATGATCCCAATATACTACGATATTCAGATG 604</p> <p>740 AGTAGKKRRRTWKKRWKRRDTRWDDADADDTARDDRRRRGDDGADGAGKKTGKRRRRD 681</p>							

QY 605 GAAAAACCCATATAATATGCTAAGTCTGGAAATTCATTTAATCAGCAA 664
 Db 680 RAYDRDAMWADAAWTTTDDDDMKRDRRRKGARRRRRTTAAWDMWTAKMDMA 621
 QY 665 TTATCAAGATGATCAACACCAATTTGGATATATATCTGAAATCCAAATTTAACAACA 724
 Db 620 KMDKTRADRDNRWADWTDAKADRDMAKARABRRDRBARADRRMTTKGKTSTA 561
 QY 725 ATCAACACCATATGAAATTTACTTGAAGTCTTGATTTAACAACAATCTCTCA 784
 Db 560 TWTWTAATAAAMWAMWMTTATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 501
 QY 785 ATTATATATAAGAGAGAGTGTGAATCTGTAGT 818
 Db 500 AATTAATAAATAAATAATTTTTTTTTTTTWT 467

RESULT 4
 CNS00EVL/c 1101 bp DNA GSS 04-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC:
 DEFINITION BACR29B23 of Rpci-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCESSION AL069706 GI:4949849
 VERSION AL069706
 KEYWORDS GSS.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 1101)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila
 melanogaster BAC library was prepared by Kazutoyo Osoegawa and
 Aaron Mamoser in Pieter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPCI-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain y2; cn bw sp, the same strain used for the BDGP's
 P1 and EST libraries. A more detailed description of the library
 and how to order individual BAC clones, the entire library, or
 filters for hybridization from the BACPAC Resource Center can be
 found at <http://bacpac.med.buffalo.edu/drosophila-bac.htm>.
 Location/Qualifiers
 source 1..1101
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone_lib="RPCI-98"
 /clone="BACR29B23"
 /note="end : T7"

BASE COUNT 419 a 91 c 60 g 299 t 232 others
 ORIGIN

Query Match 6.1%, Score 62.4; DB 121; Length 1101;
 Best Local Similarity 34.5%; Pred. No. 0.00017;
 Matches 149; Conservative 78; Mismatches 204; Indels 1; Gaps 1;

QY 581 AATATACACAGTATTCAGATGAAAAAACCCATATAAATATATGCTAGCTCG 640
 Db 1099 WWTATWTTTWTWTWTATATATATTAATTAATTTWTTATATATWTAATW 1040
 QY 641 GGAATATTCATTTAATGACAAATTTATCAAGATGATCAACACCAATTTGATATATA 700
 Db 1099 WWTATWTTTWTWTWTATATATATTAATTAATTTWTTATATATWTAATW 1040

Db 1039 ATATATWTTTWTAAATATAMWTTAATATAATATATWATAMWATATATAMATAAM 980
 QY 701 CTGAATCCCAATTTACAGCAATCAACACATATGAAATATATTTACTGAGCTT 760
 Db 979 TATWATWTAATAATWATWATWATWATWATWATWATWATWATWATWATWATWAT 920
 QY 761 GTGATTTAACAACAATCTTCACATTTATATAAAGAGAGTGTGAATCTTCTAGTCG 820
 Db 919 TTTWTTATWTTTATATWTTWAAAAAAMWATWATTTTWTWTTATWATWATWAT 860
 QY 821 AGCAGGTCCACACACTACTGACATTCATCTATATATAGATGATTTATCTCT 880
 Db 859 TAWTTTTTTTWTWTTATATWATWATWATWATWATWATWATWATWATWATWATW 801
 QY 881 ATTATGCCCAATTAATGCGGATCGCAATATATCAATTTATCAACAATGATG 940
 Db 800 WTTTATTTATWTTATWATWATWATWATWATWATWATWATWATWATWATWATWAT 741
 QY 941 TGATTGAGATACAGATGCGAACCATTTGAAATGTTCATTCGAGTTATTAATCAA 1000
 Db 740 WTTTATTTWTTTAAATWATWATWATWATWATWATWATWATWATWATWATWATW 681
 QY 1001 ATGTAAATTA 1012
 Db 680 TTATTAATTA 669

RESULT 5
 CNS0155H 1001 bp DNA GSS 26-JUL-1999
 LOCUS Drosophila melanogaster genome survey sequence SP6 end of BAC:
 DEFINITION BACN13C23 of DrosBAC library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCESSION AL105023 GI:5617037
 VERSION AL105023
 KEYWORDS GSS.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 1001)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (EDGP) -
<http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC
 library (Dros BAC) was made by Alain Billaud at CEPH (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MRC
 project grant. The DNA was prepared from embryos by Alain Bucheton
 and Genevieve Payan. It has been constructed in the vector
 pBelobAC11.
 Location/Qualifiers
 source 1..1001
 /organism="Drosophila melanogaster"
 /plasmid="pBelobAC11"
 /db_xref="taxon:7227"
 /clone_lib="DrosBAC"
 /clone="BACN13C23"
 /note="end : SP6"

BASE COUNT 266 a 219 c 134 g 150 t 232 others
 ORIGIN

Query Match 6.0%; Score 61.8; DB 121; Length 1001;
 Best Local Similarity 32.7%; Pred. No. 0.00023;
 Matches 145; Conservative 84; Mismatches 215; Indels 0; Gaps 0;

QY 353 ATGATGAACGGGATCAATATATACAGACTTTTAAAGCAAAAGCAAGCACTGC 412
 Db 1099 WWTATWTTTWTWTWTATATATATTAATTAATTTWTTATATATWTAATW 1040

Db 553 AAGRVARWMDGAGAAGADAGATATNAGAAAAAAGRGMAAABDAGRASAAREKTKV 612

QY 413 CACAATTACAGTGAAGTCTGCAAGTTAATGTTAACAAGCAATGATTAATGAC 472

Db 613 GAGGAGAAAAAGWDADDAATWRRKAAARAKAADWAAAAACAAAAVHNAHNAAGKA 672

QY 473 AAGTCAATGATTTACTAACAGAGGTTAAACAGATGCTTAAGTAAAGCATGAC 532

Db 673 AARRRAAAAAAARAGDAAAAAAAGRAARDAAAAAAATAAAMWATARRAARRDA 732

QY 533 ACAGACGAGTGTAACTGAAGACGTAAGTGAATAGATCCCAATATAGTACAC 592

Db 733 RAATAAATATATWTAAMWMAAAMWMAAATAAATTTATATTAATAAATAAAMWTA 792

QY 593 GTATTCAGATGAAAAAACCCCTATATAAGTATATGCTAGTGGCAATATTCATT 652

Db 793 AAAAATATTTTAMWMAAATAAATAAATAAATAAATAAATAAATAAATAAATTTT 852

QY 653 TTAATCAGCAATTTATCAAGATGAATCAACACCAATTTGATATAGTGAATCCAA 712

Db 853 TAAAMMAAAMWATWATATTAADAAAAAATAAATAAATAAATAAATAAATAAATTTT 912

QY 713 ATTAAACAAGCAATCAACACATATGAATTAATTTACTGGAAGTCTGTGATTTACAA 772

Db 913 AAMADMTTWTWTWTAAMWMAAATAAATWTAATWTAATWTAATWTAATWTAATWTAAT 972

QY 773 CAATCTCTCAATTTATATATAA 796

Db 973 WAAAAWMAATWTAATWTAATWTAATWTAATWTAATWTAATWTAATWTAATWTAAT 996

RESULT 6

CNS00EVL 1101 bp DNA GSS 04-JUN-1999

LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC: BAC329823 of RPCL-98 library from Drosophila melanogaster (fruit fly); genomic survey sequence.

DEFINITION

ACCESSION AL069706.1 GI:4949849

VERSION

KEYWORDS

SOURCE

ORGANISM

fruit fly.

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 1101)

AUTHORS

Genoscope.

TITLE

Direct Submission

JOURNAL

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

COMMENT

- Web : www.genoscope.cns.fr

- determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see : <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammosier in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCL-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Source

1..1101

/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/clone_lib="RPCL-98"

/clone="BACR29823"

/note="end : T7"

BASE COUNT 419 a 91 c 60 g 299 t 232 others

ORIGIN

Query Match 6.0%; Score 61.8; DB 121; Length 1101;

Best Local Similarity 31.8%; Pred. No. 0.0023;

Matches 183; Conservative 95; Mismatches 297; Indels 0; Gaps 0;

QY 371 AATTATACCAAGCTTTTAAAGCAAAAGCAAGCAACTGCCACAATTTACAGTAAG 430

Db 492 AATWAAAAAAMWMAATTTTTTMMWMTWTTWTTWTAATWTAATWTAATWTAAT 551

QY 431 TATCTCAGATTTAGATGTAACAAGCAAGATGATATGCAAGTCAATGATTTACTA 490

Db 552 TTAAMWMAATWTAATWTAATWTAATWTAATWTAATWTAATWTAATWTAATWTA 611

QY 491 ACAAGAGGTTAAACAAGATGCTATTAAGTAAAGACGACGACGAGTGTAACTG 550

Db 612 AAAATATTTTWTATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAAT 671

QY 551 GAAGACGTACGATTAATGATGATCCCAATATCTACAGTATCAAGTGAATAA 610

Db 672 ATTAATTAATTTWAAATWMAAATAAATAAATAAATAAATAAATAAATAAATAA 731

QY 611 ACCCTTAAGTAATATATGCTTAAGTCTGGAATATTTCAATTTATCAATTTATC 670

Db 732 WAATWMAAAMWMAATWMAATWMAATWMAATWMAATWMAATWMAATWMAATWMA 791

QY 671 AAGATGAATCAACACCAATTTGATATATAGTGAATAATCCAAATTTTACAGCAAT 730

Db 792 WAATWMAAAMWMAATWMAATWMAATWMAATWMAATWMAATWMAATWMAATWMA 851

QY 731 CACATATTAATATTTACTGTAAGTCTGATTTACACAAATCTTCAATTTAT 790

Db 852 AAAATTAATATWMAATWMAAATAAATAAATAAATAAATAAATAAATAAATAA 911

QY 791 ATAAAGAGAGTGTGAAGTCTGCTAGCAGCAGCTCAACCACTACTTCAAGATCT 850

Db 912 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 971

QY 851 CCATCTATATATAGATTAATTTCTATATATGCCCCGAATTAATTTGGCGATCG 910

Db 972 AATWTAATWTAATWTAATWTAATWTAATWTAATWTAATWTAATWTAATWTA 1031

QY 911 GAATATATCAATTTTCAACAAGATGATGATT 945

Db 1032 TAATATATWTAATWMAATWMAATWMAATWMAATWMAATWMAATWMAATWMA 1066

RESULT 7

CNS01807/c 1101 bp DNA GSS 26-JUL-1999

LOCUS Drosophila melanogaster genome survey sequence SP6 end of BAC: BACN3707 of DrosBAC library from Drosophila melanogaster (fruit fly); genomic survey sequence.

DEFINITION

ACCESSION AL108721.1 GI:5629025

VERSION

KEYWORDS

SOURCE

ORGANISM

fruit fly.

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 1101)

AUTHORS

Genoscope.

TITLE

Direct Submission

JOURNAL

Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

COMMENT

- Web : www.genoscope.cns.fr

- determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - <http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billand at CPH (Centre

d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.

FEATURES

Location/Qualifiers
1. .1101

Source
/organism="Drosophila melanogaster"
/plasmid="pBelobAC11"
/db_xref="taxon:7227"
/clone_11b="DrosBAC"
/clone="BACN37F07"
/note="end : 5p6"
BASE COUNT 179 a 70 c 85 g 448 t 319 others
ORIGIN

Query Match 5.9%; Score 61.2; DB 121; Length 1101;
Best Local Similarity 42.5%; Pred. No. 0.00031;
Matches 182; Conservative 21; Mismatches 221; Indels 4; Gaps 1;

QY 371 AATTACCAGACTTTTAAAGCAAAAGCAACGACCTGCCACAATTACAGTGAAG 430
DB 562 AMAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 503
QY 431 TATCTGCAAGTTAGATGTAACAAGCAATGATATGACAAAGTCAATGATTACTA 490
DB 502 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 443
QY 491 ACAAGAGGTTAACAAGATGCTCTAAGTTAGACATCGACACGCGAGTCTTACTG 550
DB 442 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 383
QY 551 GAAGCGTACGTTAATAGATGATCCACATATATCTACAGTCTCAAGTGAAGAAA 610
DB 362 AATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 323
QY 611 ACCCTAATAAGTATATGCTCTAAGTCTGGAATATTCATTTTATCAGCAATTTATC 670
DB 322 AATAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 267
QY 671 AAGATGAATCAACACCAATTGATATATCTGAAATTCCTCAATTTAACAAGCAATCAAA 730
DB 266 AATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 207
QY 731 CACATATGAATTTTACTGTAAGTCTGTGATTTAACAACAATCTCTCCCAATTTAT 790
DB 206 WAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 147
QY 791 ATAAAGA 798
DB 146 AAAAAAA 139

RESULT 8
CNS001J 1101 bp DNA GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC:
DEFINITION BACR30P10 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL068307
VERSION AL068307.1 GI:4958538
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE Genoscope.
1 (bases 1 to 1101)
AUTHORS Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 Evry cedex - FRANCE (E-mail : seqrefegenoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mammose in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Location/Qualifiers

Source
1. .1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_11b="RPCI-98"
/clone="BACR30P10"
/note="end : T7"
BASE COUNT 507 a 148 c 112 g 171 t 163 others
ORIGIN

Query Match 5.9%; Score 60.2; DB 121; Length 1101;
Best Local Similarity 37.0%; Pred. No. 0.00053;
Matches 160; Conservative 54; Mismatches 219; Indels 0; Gaps 0;

QY 369 ACAATTATACCAAGCTTTTAAAGCAAAAGCAACGACCTGCCACAATTACAGTGA 428
DB 121 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 180
QY 429 AGTATCGCAAGTTAGATGCTTAACAAGCAATGATTAATGACAAAGTCAATGATTAC 488
DB 181 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 240
QY 489 TACAAAGAGTTAACAAGATGCTATAGTTAAGCATCGACAGCAGCGATTAC 548
DB 241 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 300
QY 549 TGAAGAAGCTGACGTTGATGATGATCCACATATATCTACAGTATTCAGATGAAA 608
DB 301 WAAATCAAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 360
QY 609 AAACCTATAAAGTATATTTGCTAAGTCTGGCAATATTCATTTTAAATCAGCAATTTA 668
DB 361 WAAWAMCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 420
QY 669 TCAGATGAATCAACACCAATTGATATATCTGAAATTCACAAATTTAACAAGCAATCA 728
DB 421 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 480
QY 729 AACACATATTAATTTTACTGGAAGTCTGTGATTTAACAACAATTTCTCACAATTT 788
DB 481 AAAAGAAMCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 540
QY 789 ATATTAAGAGCA 801
DB 541 WAAAAAGACRWGA 553

RESULT 9
CNS002J 1101 bp DNA GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR05N11 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL061936
VERSION AL061936.1 GI:4940214
KEYWORDS GSS.
SOURCE fruit fly.

[illegible]

RESULT	11
CNS009DO	
LOCUS	781 bp DNA GSS 03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence TE73 end of BAC # BACR1009 of RPCT-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION	AL053444	
VERSION	AL053444.1	GI:4934889
KEYWORDS	GSS.	
SOURCE	fruit fly.	
ORGANISM	Drosophila melanogaster	

REFERENCE 1 (bases 1 to 781)

FILE
 Direct submission
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : JOURNAL

COMMENT

The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.flyritzy.org>. The BDGP *Drosophila melanogaster* BAC library was prepared by Kazuhiro Oseegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's 101 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
SOURCE

```

BASE COUNT      440 a      6 g      189 t      75 others
ORIGIN

```

Query Match	5.68;	Score 58;	DB 121;	Length 781;
Best Local Similarity	42.7%;	Pred. No. 0.0017;		
Matches 178;	Conservative 18;	Mismatches 221;	Indels 0;	Gaps 0

[illegible]

Db	177	AAAAAAAAAAAAAAAAAATTATTAATAAAAAAAAAAATTTAAATATATATTAATAA	236
Qy	500	TTAACAAGATGCTATTAAGTAAAGACATGCACACGACGAGTGTTAATCGAAGACGTA	559
Db	237	AAAAAAAAAAWTTTTTTAAAAATTAATTTTAAAAATTTAAWATTTATATTTAAAAATTTAAA	296
Qy	560	CAGTGGATTGATGATCCCAATATCTCACGATTCAGATGGAATAAACCCATTAA	619
Db	297	AAATTTAAAAAAATTAWAAAAAATAAAAAAAAAATTAWAAATTAWAAAAAATAWAAAA	356
Qy	620	AAGTATATTTGCTAGCTGGAGATATTCATTTTAATGACAAATTTATCAAGATGAAT	679
Db	357	AAAAWATAAAAAAWMTATTTATTAATAATTTATTAATTTAATTAATTAATTAATTAATTA	416
Qy	680	CACACCAATTTGATATATATCTAAAAATCCAAATTTAAACAGCAATCAACATATTG	739
Db	417	AAWTAATATATTAATAAAAAATTAATAAAAAATTTAAAAAAATTAATAAAAAATTAATTA	476
Qy	740	AAATTAATTAAGCTGTGATTTTCAACCAATTTCTCAAAATTTATATTAATAA	796
Db	477	ATAATTAATAATAAATAATTAACAAAAATAATTAATTAATAATAATTAATTAATTAATAATA	533

RESULT	12
CNS00FUH	
LOCUS	996 bp DNA
DEFINITION	GSS 03-JUN-1999
	Drosophila melanogaster genome survey sequence TET3 end of BAC:

ACCESSION	AL071063	
VERSION	AL071063.1	GI:4951105

ORGANISM *Drosophila melanogaster*

REFERENCE 1 (bases 1 to 996)

TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

COMMENT

The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.fruitfly.org>. The BDGP *Drosophila melanogaster* BAC library was prepared by Kazutoyo Oseegawa and Aaron Mammoler in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPci-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's 11 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source

```

BASE COUNT      383 a      164 c      81 g      171 t      197 others
ORIGIN
    /organism="Drosophila melanogaster"
    /db_xref="taxon:7227"
    /clone_lib="RPCI-98"
    /clone="BACR31021"
    /note="end : TET3"

```

Query Match	5.6%	Score 58;	DB 121;	Length 996;
Best Local Similarity	34.0%;	Pred. No. 0.0017;		
Matches 177:	Conservative	81;	Mismatches 260;	Indels 2; Gaps 1

FEATURES	source
LOCUS	BEA27889/c
DEFINITION	BEA27889 1218 bp mRNA EST 24-JUL-2000
ACCESSION	PSR6899 ITC PSR Barley Inflorescence Library Hordeum vulgare cDNA
VERSION	BEA27889
KEYWORDS	BEA27889.1 GI:9425732
SOURCE	EST.
ORGANISM	Barley.
REFERENCE	Hordeum vulgare
AUTHORS	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.
	1 (bases 1 to 1218)
	Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier,S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P., Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Jondrier,P., Langridge,P., Lazo,G.R., Lin,J., McGuire,P., Ogihara,Y., Pecchioni,N., Qualset,C., Schuch,W., Selvaraj,S., Shariflou,M., Sorrells,M., Warburton,M. and Wenzel,G.
TITLE	International Triticace EST Cooperative (ITC): Production of Expressed Sequence Tags for Species of the Triticace
JOURNAL	Unpublished (2000)
COMMENT	Contact: Bailey P
	Cereals Group, John Innes Centre
	Norfolk, Norwich NR4 7UH UNITED KINGDOM
	Tel: 44 1603 452571 ext. 2587
	Fax: 44 1603 502241
	Email: paul.bailey@bbsrc.ac.uk
	International Triticace EST Cooperative (ITC)
	http://wheat.pw.usda.gov/genome.
	Location/Qualifiers
	1..1218
	/organism="Hordeum vulgare"
	/cultivar="Forester"
	/db_xref="taxon:4513"

```

/cdclone="PSR6899"
/cdclone_lib="tREC PSR Barley Inflammation Library"
/rclone_type="inflammation"
/dev_stage="inflammation; development"
/note="Vector: Lambda ZAP II (Stratagene)"

BASE COUNT      201 a      160 c      158 g      530 t      169 others

ORIGIN

Query Match      5.5%; Score 56.8; DB 35; Length 1218;
Best Local Similarity 38.3%; Pred. No. 0.0033;
Matches 174; Conservative 0; Mismatches 279; Indels 1; Gaps 1;

OY 336 ATGAACGGGCATCACAAATTATACCAAGCCTTTTAAAGCAAAAGCAACATCGCCAC 415
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1166 AAAAAAGATNANNANNAANNTAAAAAANTAAANNAACCAANNAANNAANNAATNNAN 1107
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 416 AAATACAGTGAAGATATCTGCAAGTTAGATGCTAAACAGCGAATGATATGACAA- 474
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1106 AANNAANNAANNAANNAANNAANNAATTAANNANNAANNAATNGAATATAAAAACAANNAAG 1047
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 475 AGTCATGATGATTACTTAACAAGAGGTTAAACAGATGCTCTAAGTTAAGCATCGACAC 534
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1046 AATGACANNAANNAANNAANNAANNAATTAANNAANNAANNAATTAANNAANNAANNAACGAGAAA 987
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 535 GACGCAGTGTTAAGTGAAGACGTACATGTAATAGATGATCCCAATATATCTACACGT 594
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 986 TAAAAANNAAGAAACAGAGNANNNANNAACAGANNACNNANNAANNAANNAANNTTAAAAA 927
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 595 ATTCAGATGGAAGAAACCCCTATTAAGTATATGTCTAAGTCTGGGATATTCATTTT 654
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 926 AAGAGNNANNAANNAANNAANNAATTAANNAANNAANNAANNAANNAANNAAGAAAAAAC 867
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 655 AATCAGCAAAATTTATCAAGATGAATCAACACCAATTTGATATATATCAAAATCCAAT 714
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 866 NAAGAAAAAANNNANNAANNAANNAANNAANNAATTTNNAACATATGAANNAANNAANNAAC 807
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 715 TTAACACACATCAACACATATGTAATTTATTTACTGGAAGTCTTGATTTAACAACA 774
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 806 AANNAANNAANNAATTAAGATATGCAAAATATNNANNAAGAAAAANNAAGATNNANNAATTAA 747
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 775 ATTCCTCACATTTTATATTAAGAGAGACTTGGAA 808
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 746 AANAGANNAATNTTNAAGAAAAAATTANNAA 713
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
CNS016EO
LOCUS 1201 bp DNA GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
BACN1E09 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL106652
VERSION AL106652.1 GI:5623072
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscophora; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1201)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector

```

FEATURES	pBelOBAC11.	Location/Qualifiers
Source	1. .1201	
	/organism="Drosophila melanogaster"	
	/plasmid="pBelOBAC11"	
	/db_xref="taxon:7227"	
	/clone_lib="DrosBAC"	
	/clone="BACN15E09"	
	/note="end : 17"	
BASE COUNT	315 a 179 c 151 g 370 t 186 others	
ORIGIN		
Query Match	5.5%: Score 56.6: DB 121; Length 1201;	
Best Local Similarity	40.8%: Pred. No. 0.0037;	
Matches	95: Conservative 38: Mismatches 100: Indels 0: Gaps 0:	
OY	563 TTGAATAGATGATCCACAAATATACACGATTCACAGATGAGAAAAACCTATATAAG 622	
Db	943 TKTWTTKGMAHTETTKMAAAAWMAARAANAANAATTTWWRASRDAAAAAAAA 1002	
OY	623 TAATATGTCCTAAGTCGTGGCAATATTCATTTTAATCAGCAAAATTTATCAAGATGAA 682	
Db	1003 AAAATTTATATATBWRABAAATTTAAATTAATTAATTAATTAATTAATTAATTAAT 1062	
OY	683 CACCAATTTGGATATATATCTGAAATCCAAATTTTAAGCAATCAACACATATTGAA 742	
Db	1063 WATWTTWTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1122	
OY	743 TTATTTACTTGAAGCTTGTGATTTACACAACTCTTCACATTTATATAA 795	
Db	1123 TTTTWTWTTTMAWATATWATWTTAAATAATTTATTTATTTATTAATATATATWTA 1175	
RESULT 15		
CNS0039R/C		
LOCUS	CNS0039R 1101 bp DNA GSS 03-JUN-1999	
DEFINITION	Drosophila melanogaster genome survey sequence Tetr3 end of BAC #	
ACCESSION	BA008010 of Rpci-98 library from Drosophila melanogaster (fruit	
VERSION	fly), genomic survey sequence.	
KEYWORDS	AL063932	
SOURCE	AL063932.1 GI:4941789	
ORGANISM	GSS.	
REFERENCE	fruit fly	
AUTHORS	Drosophila melanogaster	
TITLE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
JOURNAL	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;	
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
	1 (bases 1 to 1101)	
COMMENT	Genoscope.	
	Direct Submission	
	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :	
	BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr	
	- Web : www.genoscope.cns.fr)	
	Determination of this BAC-end sequence was carried out as part of a	
	collaboration with the Berkeley Drosophila Genome Project (BDGP).	
	The BDGP is constructing a physical map of the Drosophila	
	melanogaster genome using these BACs. For further information	
	please see http://www.fruitfly.org The BDGP Drosophila	
	melanogaster BAC library was prepared by Kazutoyo Osoegawa and	
	Aaron Mamoser in Pieter de Jong's laboratory in the Department of	
	Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,	
	NY. The library is named Rpci-98 and was constructed by partial	
	ECORI digestion of Drosophila DNA provided by the BDGP from the	
	isogenic strain y2: cn bw sp, the same strain used for the BDGP's	
	PL and EST libraries. A more detailed description of the library	
	and how to order individual BAC clones, the entire library, or	
	filters for hybridization from the BACPAC Resource Center can be	
	found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .	
	Location/Qualifiers	
	1. .1101	
	/organism="Drosophila melanogaster"	
	/db_xref="taxon:7227"	

```

/clone_11b="RPCL-98"
/clone="BACROBIO10"
/notes="end : TET3"
BASE COUNT      170 a      130 c      84 g      469 t      248 others
ORIGIN

Query Match      5.5%: Score 56.2; DB 121; Length 1101;
Best Local Similarity 31.7%: Pred. No. 0.0045;
Matches 234; Conservative 140; Mismatches 350; Indels 14; Gaps 3

QY      69 TGTTCGTAAGTAAAGAGTATGAGCATTTGGTATGATGGTCACACCTTGAGAAAAGTGACAA 128
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1069 WKAKAGDWMAAAADAGADGAGTGTGTAGTAKAGGTDDAWGAKAATKWAGADAWAKATKAG 1010
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      129 GCATCCGAGGTTCAAGACACTGTATATGCGACACAAAATGCTGAAGTGGCAGCATTTA 188
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1009 AKATKKAAWRRKWRADWMDAATAKAAKADRDAGADAMWRATGAGTGMNNAAGTMAWRTATA 950
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      189 TATTACGTTTGA----GCCATGTACTCAATTTTGGTTCAACACACCCTGCTGTAAACAAA 244
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      949 KATTAAGATATRAAKTBTMGTRRTATATATATATATATATATATATATATATATATATATATAT 890
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      245 TTATGATTGTGAAGATAGCAAAAAGTAGATATACGCAACAAAGACAAATTCGTTAGACAC 304
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      889 TTTADADATATTTKMTATARGGGRRAKRTATKAKATKAGADTATATADKKAATKAWAD 830
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      305 ATGCGTATGAGACGCTACGGGCTCAGCGTATTTAGCGTTGAATGCGTTGATGATGAACGG 364
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      829 AMGTAKGTATATATKWDGTAKRKGGRTGAGTAAADACARAARAKRMRDAAAAAAGA 770
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      365 CATCACAATTTATACCAACACTTTTAAAGCAAAAAGCAAGCAACTGCCACAACTTACAG 424
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      769 GAGAAATATAAAAKKMAATATATGKRKRAGDAAGARGGADGKGKGGTGAAAAA 710
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      425 TGAAGATCTCGCAAGTTTAGATGCTAAACAACGCAATGATATGACACAACTCAATGA 484
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      709 AKAAAKTDTJAWMAAAATATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 650
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      485 TTACTAACAAAGAGGTTAAACAGATGCTTAAGTTAAGATCATGACACGACGACAGTGT 544
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      649 WKAKKAMDAKRAATATAAAAAAAG---AKTAKKMAAAAAAADAADAAAAAATAAAAAA 593
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      545 TAACTGAAGAAGCGTACAGTTGAATTAGATGCACAAATATCTACAGTATTCAGATG 604
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      592 AAAAAAGAAARGTAWGWTGAKNCGAAAAAATAAAAAAATAAAAAAATAAAAAAATA 533
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      605 GAAAAAACCCCTATATAAATAATATGTTAGTCTGGGAATATTCATTTTAAATCAGCAA 664
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      532 AAAAAAAGKRAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 473
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      665 TTATACAGATGATGATCA--CACCATTTGGATATATATCTGAAATGCCAATTTA 717
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      412 WTMAAAMMAAAMMAAAMMAATMTGMAAAATMTAATTTAATMAAATMAAAMMAAARTAAA 413
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      718 ACAAGCATCAAAACATATTTGAATTTATCTTGAAGTCTGTGATTTAAACAACAT 777
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      412 AAAAAAAMMAAAMMAAAMMAATMTAARATGTGATCTAATHACHYAAATTCYMY 353
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      778 CTTCACAATTTATATAA 795
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      352 YMYCCMTYTCMCMAA 335
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Search completed: November 19, 2000, 03:36:00
Job time: 12209 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 04:23:24 ; Search time 2632.38 seconds
(without alignments)
2105.493 million cell updates/sec

Title: US-08-978-456-3
Perfect score: 1269
Sequence: 1 AANCACCAATCCNATTGGGA.....GTTAAACAAGATCTATAG 1269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1033670 seqs, 2183789903 residues

Total number of hits satisfying chosen parameters: 2067340

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_p11:*
8: gb_p12:*
9: gb_p11:*
10: gb_p12:*
11: gb_p13:*
12: em_fun:*
13: em_hum1:*
14: em_hum2:*
15: em_in:*
16: em_om:*
17: em_or:*
18: em_ov:*
19: em_pat:*
20: em_ph:*
21: em_pl:*
22: em_ro:*
23: em_sts:*
24: em_sy:*
25: em_un:*
26: em_v1:*
27: gb_htg1:*
28: gb_htg2:*
29: gb_in1:*
30: gb_in2:*
31: em_ba1:*
32: em_ba2:*
33: em_hum3:*
34: em_hum4:*
35: gb_pr4:*
36: gb_htg3:*
37: gb_htg4:*
38: gb_htg5:*
39: gb_htg6:*
40: gb_htg7:*
41: em_htg1:*
42: em_htg2:*
43: em_htg3:*

44: em_hum5:*
45: gb_p13:*
46: gb_p15:*
47: gb_htg8:*
48: gb_htg9:*
49: gb_htg10:*
50: gb_htg11:*
51: gb_htg12:*
52: gb_htg13:*
53: gb_htg14:*
54: gb_in3:*
55: gb_htg15:*
56: gb_htg16:*
57: gb_htg17:*
58: em_htg4:*
59: em_htg5:*
60: em_htg6:*
61: em_htg7:*
62: em_hum6:*
63: gb_htg18:*
64: gb_htg19:*
65: gb_ba3:*
66: em_htg8:*
67: em_htg9:*
68: em_htg10:*
69: em_htg11:*
70: em_htg12:*
71: em_htg13:*
72: em_htg14:*
73: em_htg15:*
74: em_htg16:*
75: em_htg17:*
76: em_htg18:*
77: em_htg19:*
78: em_htg20:*
79: em_htg21:*
80: em_htg22:*
81: em_htg23:*
82: gb_p16:*
83: gb_p17:*
84: gb_htg20:*
85: gb_htg21:*
86: gb_htg22:*
87: gb_htg23:*
88: gb_ro:*
89: gb_sts1:*
90: gb_sts2:*
91: gb_sy:*
92: gb_un:*
93: gb_v11:*
94: gb_v12:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	
1	352.2	27.8	3336	5	AR069579	AR069579 Sequence	
2	289.2	22.8	3227	1	AF269646	AF269646 Staphyloc	
3	289.2	22.8	3618	1	AF269345	AF269345 Staphyloc	
4	288	22.7	3721	1	AF269926	AF269926 Staphyloc	
5	176	13.9	2902	1	AF270130	AF270130 Staphyloc	
6	103.6	8.2	6006	2	BSRIB	X51510 B. subtilis	
7	103	103	8.1	28206	1	BACDIA	L09328 Bacillus su
8	103	103	8.1	218470	2	BSUB0013	Z99116 Bacillus su
9	100.8	7.9	4286	2	BARIBGENS	X95955 B.amyloliqu	
10	98.8	7.8	18920	57	AC074319	AC074319 Staphyloc	
11	84.8	6.7	14364	1	AE000675	AE000675 Aquifex a	
12	69.4	5.5	18073	65	U32775	U32775 Haemophilus	

C 13 68.4 5.4 11695 1 AE004298
14 67.6 5.3 2949 2 ECNUSB
15 67.6 5.3 11692 2 AE000148
16 67.6 5.3 139818 2 EC082664
17 64.4 5.1 2759 1 AF002857
18 64.4 5.1 4312 1 AF027202
19 63.8 5.0 14376 1 AE002531
20 62.6 4.9 11544 1 AE001343
21 60.6 4.8 349061 2 NMA222491
22 60.2 4.7 11404 1 AE003934
23 57 4.5 2307 2 BHE132928
24 56.4 4.4 105795 65 SYCCFNC
25 53.8 4.2 11456 1 AE002277
26 53.6 4.2 24887 2 BACSERA
27 52.6 4.1 3332 1 AF269888
28 52.6 4.1 170227 48 AF269712
29 52.4 4.1 170227 48 AF023138
30 52.2 4.1 110000 87 PFMA14P1_1
31 52 4.1 164353 37 AC011978
32 51.8 4.1 170413 64 AL159155
33 51.6 4.1 145670 10 AC008132
34 51.2 4.0 189841 64 AL139811
35 51 4.0 105567 35 AP001069
36 51 4.0 340000 35 AP001708
37 50.8 4.0 3542 45 YSCWTCG16
38 50.6 4.0 163226 27 AC004157
39 50.6 4.0 167712 84 AL356794
40 50.4 4.0 158482 50 AC025348
41 50.2 4.0 302250 83 U82671
42 50 3.9 1158 30 AF044857
43 50 3.9 184207 64 AL353748
44 49.8 3.9 1867 45 MTSCMJ23
45 49.8 3.9 12811 1 AE001819

ALIGNMENTS

RESULT 1
LOCUS AR069579 3336 bp DNA
DEFINITION Sequence 7 from patent US 5891672.
ACCESSION AR069579
VERSION AR069579.1 GI:7220467
KEYWORDS
SOURCE .
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 3336)
AUTHORS Wang,M., Ward,J.M., Warren,R.,Lloyd,Nicholas,R.,Oakley,
Palmer,L.,Marie, Pratt,J.M., Knowles,D.,JustinCharles,
Loretto,M.,Arthur, Mooney,J., Black,M.,Terence,
Burnham,M.,KarlRussell, Debouch,C., Fedon,J.,Craig, Hodgson,J.,Edward,
Jaworski,D.,Dee, Reichard,R.,Winfield, Rosenbery,M., Trainl,C.,Michael
and Zhong,Y.Y.I.
TITLE Polynucleotides encoding GTP cyclohydrolase II (riba)
JOURNAL Patent: US 5891672-A 7 06-APR-1999;
FEATURES Location/Qualifiers
source 1..3336
BASE COUNT 1202 a 459 c 678 g 997 t
ORIGIN

Query Match 27.8%; Score 352.2; DB 5; Length 3336;
Best Local Similarity 88.4%; Pred. No. 8.6e-56;
Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;

OY 734 ATGATTTGCATTCACCTTCCAAATATGTCACAGTCACAGTCCTTATACCAACC 793
Db 1 ATGGATTATGCGATTCATTCCAATATGTCACAGTCACAGTCCTTATACCAACC 60
OY 794 GTTGCGCTGTTGTAGTAATGAAGTAGAGATTGTTGATTGTGACACTTGAGAAA 853

Db 61 GTTGCGCTGTTGTAGTAATGAAGTAGAGATTGTTGATTGTGACACTTGAGAAA 120
OY 854 GGTGACAGCATCGCGAGGTCACAGCACTTATATGACACAAANAAATGCGAAGTGCG 913
Db 121 GGTGACAGCATCGCGAGGTCACAGCACTTATATGACACAAANAAATGCTGAAGTGCG 180
OY 914 ACAGTTATATAGCTTACAGTCATAGTCATTTGGTTACACCAACCCTGTGTAAAC 973
Db 181 ACAGTTATATAGCTTACAGTCATAGTCATTTGGTTACACCAACCCTGTGTAAAC 240
OY 974 AAAATTATGATTTGATAGATAGCAAAAGTACTATTACCAACANAGCAATTCCTTAG 1033
Db 241 AAAATTATGATTTGATAGATAGCAAAAGTACTATTACCAACANAGCAATTCCTTAG 298
OY 1034 ACACAGCTGGGATGATGACCTTACGGGGCTCCACGGGATTTGAGGGTGAATGCGTTG 1093
Db 299 ACACAGCTGGGATGATGACCTTACGGGGCTCCACGGGATTTGAGGGTGAATGCGTTG 351
OY 1094 GATGATGACGCGCATCACAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCA 1153
Db 352 GATGATGACGCGCATCACAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCA 408
OY 1154 CTGCGCAAAATTTACAGTGAAGTCTTGAAGTTAGTGGTAAACAAACGAATT 1213
Db 409 CT--GCCCAAAATTTACAGTGAAGTCTTGAAGTGAAGTGAAGCAAGCG--AAT 462
OY 1214 GATATGACAAAGTCATGATTTACTACAAAGGTTAAACAAAGATGTCTATA 1268
Db 463 GATATGACAAAGTCATGATTTACTACAAAGGTTAAACAAAGATGTCTATA 517

RESULT 2
LOCUS AF269646/c 3227 bp DNA BCT 01-AUG-2000
DEFINITION Staphylococcus epidermidis strain SRI clone step.1015d10 genomic
sequence.
ACCESSION AF269646
VERSION AF269646.1 GI:9623542
KEYWORDS
SOURCE .
ORGANISM Staphylococcus epidermidis.
REFERENCE Staphylococcus epidermidis.
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Staphylococcus.
1 (bases 1 to 3227)
AUTHORS Kimmery,M.J., Taylor,J.,David, Nelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenbee,S., Ashanti,C., Altshuler,G., Mamou,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Transposon-mediated sequencing of the staphylococcus epidermidis
genome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3227)
AUTHORS Taylor,J.,David, Kimmery,M.J., Nelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenbee,S., Ashanti,C., Altshuler,G., Mamou,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2000) Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES Location/Qualifiers
source 1..3227
BASE COUNT 980 a 599 c 454 g 1194 t
ORIGIN

Query Match	22.8%	Score 289.2	DB 1:	Length 3227:
Best Local Similarity	69.4%	Pred. No. 3.7e-44:		
Matches 481:	Conservative 0:	Mismatches 171:	Indels 41:	Gaps
QY 346	AGATATATTTAAATTAACCGTCATTTATGTTTTTTAGAAACATATAGTATCATTT 405			
DB 2327	AAAATCTTTAATAAGTAAAGTAAATATGTAATTTTAAATTTAAATATAGTAGATTT 2268			
QY 406	TAAATGTAGTTGACACTACTAGTACTCAATATATCTATATACCAATTTCAATATATCTT 465			
DB 2267	ACAAATGTAGTGAAGAACTACTGACTCATATTAAT-ATAACAAATTTCAATATAT-TTCTT 2211			
QY 466	TCGGGGCAGGGGGAATTCACACCGGAGTAGTAATAAAGCCGACCGTGGCAATAT--T 522			
DB 2210	TCGGGGCAGGGGGAATTCACACCGGAGTAGTAATATGAGCTTGGCCACCGACCATATTTGCT 2151			
QY 523	GTTTCATATATAGTGGCTATCTAGTAGATTTCTAGAGCTTTAGAGCGCAGTAAAGTCTGGATGGG 582			
DB 2150	TTTTCATATGTATGGCTGCTATCTAGTGTGAATCTAAGCCGACAGTATAGAGTGTGGATGGG 2091			
QY 563	AGAAAGAAATGTTAATTAATTCGACAAAGATATAGTAGCGTATTTGTAAAA----- 630			
DB 2090	AGAAAGAAATGTTAATTAATTCGACAAATTCGTAATTCGTAATATGATATAGCGTATTTGAT 2031			
QY 631	-----ATGTGACAAATAGCGCTTTTAAAGCATAAATTTTTCCTCTTG 674			
DB 2030	ACTTATGTGCAATTTTAAAGTATGATTAATAGCGCTTATTTAAAGATATTAATATCTCTTAT 1971			
QY 675	CATTTAATTCATGATGTGAGGATTTTTTGTATTAGAGGTGATCATTTGATGATCAATTTA 734			
DB 1970	GCCATCAATATATAGTATTAAGGTTTTTT-----GAGGTGATCAATTTGATGATCAATTTA 1918			
QY 735	TGATTTATCGATTCACCTCCAAATATGTGACAAAGTTCANACAGGTGTTAATCCACCGG 794			
DB 1917	TGAGATGATGCTATTCACATAGCAAAAATGCTAAATGACAAACAGGTGTTAATCCACCGG 1858			
QY 795	TTGGCGCTGTTGACTTATGAAGGTAGATGTTGTTGATTTGCTGCACACTGGAANAAG 854			
DB 1857	TAGGATCCGTTGTTGTAAAAACGGTAGGATTTGATGTTAGTGCACATTTTAAAAAAGG 1798			
QY 855	GTCGACAAACATCGGAGGCTTCACAGCACTGATATGTCGACAAACANAATGCTGAGGTGGCA 914			
DB 1797	GAGATTAACATGCCAAGTACAGCTATTTGAATGTGGCAGGTTTAAATGCCAACAGGTGCTA 1738			
QY 915	CGATTTATATTAAGTTAGAGCCCATGTAGTCAATTTTGGTTCAACACCACCTGTGTAAACA 974			
DB 1737	CCATATACGTTTCATTTAGAACCTTCACACACCATGTTCAACACACCTGTGTGCATATA 1678			
QY 975	AAATATTGATTTGTAAGTAGCANAAATGAT 1007			
DB 1677	AAATCATTTGAAGCGGCAATATTAAGGTCATCT 1645			
RESULT 3				
AF269345/c	3618 bp DNA	BCT	03-AUG-2000	
LOCUS				
DEFINITION	Staphylococcus epidermidis strain sri clone step.1002d07 genomic			
ACCESSION	AF269345			
KEYWORDS	AF269345.1 GI:9664666			
ORGANISM	Staphylococcus epidermidis.			
SOURCE	Staphylococcus epidermidis			
REFERENCE	Bacteria; Firmicutes; Bacillus/Clostridium group;			
AUTHORS	Bacillus/Staphylococcus group; Staphylococcus.			
	1 (bases 1 to 3618)			
	Kimmel, W.J., Taylor, J. David, Neilsen, A.J., Godlewski, M.M.,			
	Rubino, M.A., Nelson, F.J., Rivers, P.R., Tortorella-Miller, I.,			
	Listenbee, S., Ashanti, C., Althuler, G., Mamo, L., Shepherd, N.S.,			
	Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and			
	Furdon, P.J.			
TITLE	Transposon-mediated sequencing of the Staphylococcus epidermidis			
	genome			

JOURNAL	REFERENCE	2 (bases 1 to 3618)
AUTHORS		Taylor, J. David, Kimmerly, W. J., Nelsen, A. J., Godlewski, M. M., Rudino, M. A., Nelson, F. J., Rivers, P. R., Torrella-Miller, I., Listebbee, S., Ashanti, C., Alshiller, G., Mamo, L., Shepherd, N. S., Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D. H., Miller, G. S. and Furdon, P. J.
TITLE	Direct Submission	
JOURNAL	Submitted (19-MAY-2000) Departments of Genomic Sciences and Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore Drive, Research Triangle Park, North Carolina 27709-3398, USA	
FEATURES	Location/Qualifiers	
SOURCE	1. 3618 /organism="Staphylococcus epidermidis" /strain="SRI" /db_xref="taxon:1282" /clone="step.1002d07"	
BASE COUNT	1099 a 709 c 518 g 1292 t	
ORIGIN		
Query Match	22.8%;	Score 289.2; DB 1; Length 3618;
Best Local Similarity	69.4%;	Pred. No. 3.6e-44;
Matches	481; Conservative	0; Mismatches 171; Indels 41; Gaps 5;
QY	346	AGATTATTTAAATTAACCGCTCATTTATTTGTTTTTAAACAAACATATAGTATCATTT 405
Db	3110	AAATCTTTTATTAAGTAACCTATTAATATTTGATTTTAAATTAATATGTAGATTT 3051
OY	406	TAAATGTAGTACATCTACTACTCAATTAATCTATTAACATTTTCATATATATCTT 465
Db	3050	ACAATGTATGAAGAAGTACACTGACTCATATTAATC-ATAACAATTTTCATATAT--TTCTT 2994
OY	466	TCGGGCGAGGCTGAATTTCCACCGCGAGTAATTAAGCTCGCGACTGCTAATA--T 522
Db	2993	TCGGGCGAGGCTGAATTTCCACCGCGAGTAATTAAGCTCGCGACTGCTAATAATGT 2934
OY	523	GTTTCATTTAGTGGCTGATCTAGTGAATTCCTAGAGCCGACATTAAGCTCGATGG 582
Db	2933	TTTGTATGTTGTGGCTGATCTAGTGTCAATCTAGAGCCGACAGTAAAGTCTGGATGG 2874
OY	583	AGAAAGATGTTAATTTATCGACAAGATATAGTACGATTTGTAAAA----- 630
Db	2873	AGAAAGAGTGAATATATACGACTTACACAATTTGCTGTAGTTATGATAGCGTATTCAT 2814
OY	631	-----ATGTGCAAAATAGGCTTATTTAACGATAAATTTTCTCCCTTG 674
Db	2813	ACTTATGTGCAATTTTAAAGTATGCTAATAGGCTTATTAAGATATATAGATATTCCTTAT 2754
OY	675	CATCTTAATTCATGATGTGAAGATTTTGTGTTATAGAGTGATCATTTTGTGTCATTTA 734
Db	2753	GCTTACATATAGGTATAGGATTTT-----GAGGTGATCAATGTGTGATTTTA 2701
OY	735	TGATTAATGCAATCAATCTCCAAATATAGTACAAAGTCANAAAGGTGTTATATCCACCG 794
Db	2700	TGATGATGCTTATCACTACATAGCAAAATAGGTAAATGAGCAACAGGTGTTATATCCACCG 2641
OY	795	TTGCGCGTGTGTGATATGAAGGTAGATTTGTTGATTTGTTGTCACACTTGAAGAAAG 854
Db	2640	TAGGATCCGTTGTGTTTAAACCGTAGATTTGTAAGTTTATAGGTGACATTTAAAAAAGG 2581
OY	855	GTGACAAAGCATGCGAGGTTCAAGCACTTGATATGCGCAACAAANATGCTGAAGGTGCGA 914
Db	2580	GAGATAAACATGCGAGGATACAGCTATTTGAAATGCGGAGTTTAAATCCCAAGGTGCTA 2521
OY	915	CGATTTAATTAACGTTGAGCGCATGTAGTCAATTTGGTTTAAACACCCGCTGTAAACA 974
Db	2520	CCATATAGCTTTCATATGAACCTTGACACACCATATGGTTCAACACCACTTGTTGTGATCA 2461
OY	975	AAATTAATGATTTAGATAGCANAAGTAACT 1007
Db	2460	AAATCATTTGAAGCGGCATATCTAAGGTCACT 2428

[illegible]

OY	735	TGGATTATGCATTCACATCTCCAAATAATAGTACAAGGTCAACAGGTTATATCCACC	G	794
Dd	238	TGGATGATGCATTAATCACTAGCAGAAAAAATGTAATTGACAAACAGGTTATATCCACC	A	199
OY	795	TTGGCGCTGTTGTAGTATATAGAAGGTAGCATTTGTTGTTATTTGGTGCACACTT	A	854
Dd	198	TAGGATCCGTTGTGTTTAAAACGGTAGAGTTGATTTGAGTGTGCACATTTAAAAAG	G	139
OY	855	GTGACAAACATGCCGAGGTTCCAGACACTGATPATGGCCACACANNAATGCTGAAGTGG	A	914
Dd	138	GAGATTAACATGCGCACTACAGACTATTGAAATGGCAGGTTTAAATATCCCAAGTGCT	A	79
OY	915	CGATTATATATACGTTATAGAGCCATGTAAGTACTATTTGGTTTCAACACACCCTGTGTA	A	974
Dd	78	CCATATAGCTTTTCATTAACACTTGACACACACATGCTTCAACACACACTGTGTGCATA	A	19
OY	975	AAATTATTGA 984		
Dd	18	AAATCATTTGA 9		
RESULT	5			
LOCUS	AF270130	2902 bp	DNA	BCT 01-AUG-2000
DEFINITION	Staphylococcus epidermidis strain sri clone step.1051a10 genomic sequence.			
ACCESSION	AF270130			
VERSION	AF270130.1	GI:9624037		
KEYWORDS				
SOURCE	Staphylococcus epidermidis.			
ORGANISM	Staphylococcus epidermidis.			
REFERENCE	1 (bases 1 to 2902)			
AUTHORS	Kimmerly,W.U., Taylor,J.David, Nelsen,A.J., Godlewski,M.M., Rudino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I., Listebsee,S., Ashanti,C., Alshuller,G., Mamo,L., Shepherd,N.S., Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and Furdon,P.J.			
TITLE	Transposon-mediated sequencing of the Staphylococcus epidermidis genome			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 2902)			
AUTHORS	Taylor,J.David, Kimmerly,W.J., Nelsen,A.J., Godlewski,M.M., Rudino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I., Listebsee,S., Ashanti,C., Alshuller,G., Mamo,L., Shepherd,N.S., Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and Furdon,P.J.			
TITLE	Direct Submission			
JOURNAL	Submitted (22-MAY-2000) Departments of Genomic Sciences and Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore Drive, Research Triangle Park, North Carolina 27709-3398, USA			
FEATURES	location/Qualifiers			
SOURCE	1..2902			
	/organism="Staphylococcus epidermidis"			
	/strain="SRI"			
	/db_xref="taxon:1282"			
	/clone="step.1051a10"			
BASE COUNT	1009 a 382 c 507 g 1004 t			
ORIGIN				
Query Match	13.9%;	Score 176;	DB 1:	Length 2902;
Best Local Similarity	68.2%;	Pred. No. 2,9e-23;		
Matches 242;	Conservative 0;	Mismatches 113;	Indels 0;	Gaps 0;
OY	653	AACGATAAATTTTTCTCCCTTGACATCTTAATTCATGATGTAGAGGATTTTGTATTATAGA	712	
Dd	2139	AGCGATGACTTTTCTACGACATGATATATGCCCTCAATATATAGTATAGGTTTTTTTGA	2198	
OY	713	GGTATCATTTTGAGTCATTTATGATATATCCGATTCATTCACAAATATGTTACACAGT	772	

[illegible]

BASE COUNT	1891 a	1167 c	1401 g	1547 t
ORIGIN				
Query Match	8.2%;	Score 103.6;	DB 2;	Length 6006;
Best Local Similarity	51.3%;	Pred. No. 6.1e-10;		
Matches	345;	Conservative	0;	Mismatches 312; Indels 16; Gaps 4;
OY	346	AGATTATATTAATTAACACCGTCATTAATGTTTTTTAGAAAACATATAGTATCATTT	405	
Db	823	ACATACCTTCCGATCAGAGGCGTATGTTTCTTCCAAATGTAATGTTATTTATTCAT	882	
OY	406	TAAATGTTAGTACATACAGTACCAATATATATACATTTCAATATATATAT	462	
Db	883	TCCGACTTTAAAAAGATCGCTATATATACCAATATAGACAAATGATTAAGATTGT	942	
OY	463	CTTTGGGGGAGGTGTAATTCACAACCGCAGTAATTAAGCCTGCGACTGTAATAT	522	
Db	943	CCTTGGGGGAGGTGTAATTCACAACCGCAGTATAGACAAATTTCTTAGAGCCC	1002	
OY	523	CTTTATATATTAATGCTGATCTAGTACATTT-----CTAGAGCCGACAGTTAAAT	573	
Db	1003	CTGACCCGCTGCTCATAGACGCGGCTGATTCAGTTTAAGCTGAAGCCGACAGTAAAT	1062	
OY	574	CTGATGGGAGAGAAAGATGTTAT--TATGACAAGATTAATGATGCTATTTGTAATA	630	
Db	1063	CTGATGGGAGAGAGATGATGATGAGCCGCTATGCAAAATGTTAAAAATGCATGTTAT	1122	
OY	631	ATGTGTACAAATAGCGTTATTTAAGCATTAATTTTCCTTGTGATCTTAATGATGAT	690	
Db	1123	TCCATATGCTAAATATACCTAAAGCCCAATTTTATTAATTTGGGGCTTTTGGACG	1182	
OY	691	GTGAGGATTTTGTGTTTATAGAGTGATCATTTGAGTCAATT-TATGATTTAGCATTC	749	
Db	1183	CTAAATTAACAAAAGAGGGGAGGAAACAAATGGAAGATATATATGAAGCTGGCCTTAG	1242	
OY	750	AACCTCCAAATATGTGTACAGGTCANACAGGTGTTATGCACCCGTTGGCGCTGTGTAG	809	
Db	1243	ATCTTCCGAGAGGCGGAGAGACAGCCGAATCCAAATCCGCTCGTGGCGCTGTGTGTG	1302	
OY	810	TTAAAGAAAGGTAGATTTGTTGTTGTCACACTTGGAAAAGAGTGACAAGCATGCGG	869	
Db	1303	TAAAGAGCGGACAAATGTGCGAATGGGCGCCCATTTAAATATGTTGAAGCTCATGACG	1362	
OY	870	AGCTTCAAGCACTGTGATATGCGACAAACAAATGCTGAAGTGGCAGCATTTATTTAGCT	929	
Db	1363	AAGTTATGTCCTCATATATGCTGAGCAGATGCAAGAGGTCGCCGATTTTCATTACAC	1422	
OY	930	TGAGGCAATGTGATCATTTTGGTTCACACACACCCCTGTGTTAAACAAATTTATGATTGA	989	
Db	1423	TGGAACCGGCGACCCATTGGAAGAAACACGCCCATGTGCAGATTTGATTACAACTCTG	1482	
OY	990	AGATAGCANAACT 1002		
Db	1483	GTATCAAAAAGACT 1495		
RESULT 7				
LOCUS	BACDIA	28206 bp	DNA	BC1
DEFINITION	Bacillus subtilis spova to sera region.			
ACCESSION	L09228			26-MAY-1995
VERSION	L09228.1	GI:410114		
KEYWORDS	3-dehydroquinolate dehydratase; aroC gene; diaminiopimelate decarboxylase; lysA gene; penicillin-binding protein; peptidyl-prolyl isomerase; phosphoglycerate dehydrogenase; ppib gene; response regulator; response regulator kinase; ribA gene; ribB gene; ribD gene; ribG gene; ribH gene; ribT gene; riboflavin biosynthesis operon; sera gene; signal peptidase; sips gene; spoA gene; spovAF gene.			
SOURCE	Bacillus subtilis (strain 168, sub-species Marburg) DNA.			
ORGANISM	Bacteria; Firmicutes; Bacillus/Clostridium group;			

REFERENCE	Bacillus/Staphylococcus group; Bacillus.
AUTHORS	1 (bases 1 to 28206)
TITLE	Yamamoto,J., Shimizu,M. and Yamane,K. Molecular cloning and analysis of nucleotide sequence of the Bacillus subtilis lysA gene region using B. subtilis phage vectors and a multi-copy plasmid, pub110
JOURNAL	Agric. Biol. Chem. 55 (6), 1615-1626 (1991)
MEDLINE	91345841
AUTHORS	2 (bases 1 to 28206)
TITLE	Buchanan,C.E. and Ling,M.L. Isolation and sequence analysis of dacB, which encodes a sporulation-specific penicillin-binding protein in Bacillus subtilis
JOURNAL	J. Bacteriol. 174 (6), 1717-1725 (1992)
MEDLINE	92193254
REFERENCE	3 (bases 1 to 28206)
AUTHORS	Sorokin,A., Zumslein,E., Azevedo,V., Ehrlich,S.D. and Serrif,P. The organization of the Bacillus subtilis 168 chromosome region between the spoVA and sera genetic loci, based on sequence data
JOURNAL	Mol. Microbiol. 10 (2), 385-395 (1993)
MEDLINE	95020538
FEATURES	Location/Qualifiers
source	1..28206
organism="Bacillus subtilis"	
strain="168"	
sub-species="Marburg"	
db_xref="taxon:1423"	
gene	1..1239
gene="spovAF"	
codon_start=1	
transl_table=1	
protein_id="AA67472.1"	
db_xref="GI:410115"	
translation="VDIVENRLNAQYERKYLDETDDYLSGLVAVIYEGAFALI DVRSPGRNEEPTEKRVGARGDFEYNTALRLRIIDRELRKMTKYGERSK TDLSICYIEDIADPDIVEIYEKEIASIDVEMADKREYFNOSYNFPLVRYTE RPDYANHVYEGHYIITVDTSPIYITPTTLFFHHVQHAEXYRAPSVCPTLRVPRG ILASTLEPLTFWLFVOPDLDPDMKRTIGAKOTHPILITITLADIGITFLMAIHT CPHTRCFISYKRRHRLYSANYAMASIKSLQTPYLPFLIPFGKALMOVLVBTAPBG AKVPSIVHPRKNRLQPTNS"	
conflict	1158..1164
gene="spovAF"	
citation=[1]	
1245..1250	
gene="lysA"	
note="putative"	
1245..2664	
gene="lysA"	
1266..1272	
gene="lysA"	
note="putative"	
1345..2664	
gene="lysA"	
codon_start=1	
transl_table=1	
product="diaminiopimelate decarboxylase"	
protein_id="AA67473.1"	
db_xref="GI:410117"	
translation="MFLGSRONOHGLEIGGDVALYLAERYCTPLVYVVALIRER AKSFKAQFISGLKAOVAYSKAFSSAMQOLAEEBLSLDYVSGGELYTAAGAPPA ERHIFHNNNSREFLRALHLRIGCIIVDNFYEIALLELDCKETGHSIDVLLITTEYV EAAHHDITIGQEDSKRGFDLHNGOTERALEOVLOSHEIQLGVHCHIGSQIDTQAF VLAERKFKKLEWDRODSYFSKVLNIGGFGIRYEDDEPLHATEVEKIIIAVENV ASRGEPIPEIWEIPGRSLVGDGTTITVGSQKEYGVQYAVADGMDNTRPALY QAKYEAAANRIGEAHDKTVSINGKCESEMDLIMPIDLPEVREGDLANFCGANGY SMANVRIIPRAVPEVNEAHLVYKRETYEDIVKLDLPFKGVKQ"	
conflict	1610..1616
gene="lysA"	
citation=[1]	
1751..1757	
conflict	

```
/gene="lysa"
/citation=[1]
2496..2502
/gene="lysa"
/citation=[1]
2504..2510
/gene="lysa"
/citation=[1]
2666..2694
complement(2666..2694)
/note="ORFX19"
complement(2693..3565)
/note="ORFX19"
/codon_start=1
/transl_table=11
/protein_id="AA67474.1"
/db_xref="GI:410118"
/translation="MKIWMGMLAAVLLMVPKVSADAAVDYIVTLGADLSESDK
OKLDEMVPDNAVTVTNKEHEYLKGIYNAIGSPAISSSIITAKKSGINVE
THNIGITDEMYLNLMTAGVADKAVYVAPEYVGTALTLKAYEVSDEAISD
VKQVANQELVTTSELDKIGNENAAALLAKIIEPRKKNQVPRNKADIEQVDDASDL
NVTLDSONDLVSLFNKMKKNADIDMGVSDLDKAKDKITIKFIESDEKNTQKVID
FEVSIWNAIVSLFK"
/citation=[1]
3347..3353
/citation=[1]
3413..3419
/citation=[1]
3428..3434
/citation=[1]
3503..3509
/citation=[1]
3819..4290
/gene="ppib"
3819..4250
/gene="ppib"
/codon_start=1
/transl_table=11
/product="peptidyl-prolyl isomerase"
/protein_id="AA67475.1"
/db_xref="GI:410119"
/translation="MKTGYFLLEDGNKIEFELYPEAAPGTANFEKLANEGFYDGLTF
HRIYGFVSOGCPHGTGTGPGYTICKETEGNPHHEGALSMHAGKDTGSGOFFI
VHEPQHLNGVHTVFGKVTGSLGELFAKKNSGDMKMEVREVG"
3982..3988
/gene="ppib"
/citation=[1]
4261..4290
/gene="ppib"
4287..4320
/note="inverted repeat"
5264..5467
/note="ORFX1"
/codon_start=1
/transl_table=11
/protein_id="AA67476.1"
/db_xref="GI:410120"
/translation="MOVLTLLEOYGYAVLVLMELIALIGPEFAGVLYVOEKLN
WVLSMLAAGLGCSEHVVLDV"
5445..5834
/note="ORFX2"
/codon_start=1
/transl_table=11
/protein_id="AA67477.1"
/db_xref="GI:410121"
/translation="MSYWIGYRGLGAPPAIOKAFASRGVLLAVIVLTFEFTVFGODRKI
ELSPLEFLVIGGELVEGIGRLEFHSLHPVHAPLESIVLVPPSEGLTFAVLFGCAY
LTVRSHKTRIQVAFFCLSLFYSSSG"
6022..6027
/gene="sips"
/note="putative"
6022..7323
/gene="sips"
6045..6051
```

```
/gene="sips"
/note="putative"
6683..6688
/gene="sips"
/note="putative"
6705..6710
/gene="sips"
/note="putative"
6740..7294
/gene="sips"
/codon_start=1
/transl_table=11
/product="signal peptidase"
/protein_id="AA67478.1"
/db_xref="GI:410122"
/translation="MKSNSVSKKSLLEMAKIYIAVIALIRNFIFAPYVDGDSM
IPTLHNRERFVNMTKITGEFDKSDIYLVGDDVHYKRIIGLPDIVEKNDQLYI
NGKIVDEPYLANAKRAKODGFHDUTDFGVKVPDNKTFVMDGNRRNSMDSRNGLGL
FTKQIAQTSFEVFPENMRKTN"
7299..7323
/gene="sips"
7388..7393
/note="putative; ORFX4"
7410..7416
/note="putative; ORFX4"
7529..7873
/note="ORFX4"
/codon_start=1
/transl_table=11
/protein_id="AA67479.1"
/db_xref="GI:410123"
/translation="MGRIKRTITLLVLLLAGGYMYINDIELKDVPFTAIGQLSSE
EEEYTIQEVKVTKIDGSEYHGAENGTIIFNGKKLNQDLDIEGDKIKAYFSKSR
IDGLIKVAKVMD"
7872..7911
/note="ORFX4"
7937..7942
7960..7965
8126..8137
/note="putative"
```

Query Match 8.1%; Score 103; DB 1; Length 28206;
Best Local Similarity 60.6%; Pred. No. 6.3e-10;

Matches 166; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

```
QY 729 AATTATGATTAATGCGATTCACCTTCCAAATATGTTACAAAGTCANRACGGTGTATATC 788
      |||||  ||  |||||  ||  |||||  |||||  |||||  |||||
DB 8276 AATTATGAAGCTGCGCTTAGATCTTGCGAAGCAGGCGAAGACAGACCAATCCATC 8335
      |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 789 CACCGTGGCGCTGTAGTATTAATGAAGTAGATGTTGTATTTGGTGCACCTTGA 848
      |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 8336 CGCTCGCGCGCTGTTCGTAAAGACGACGAATTTGCGAATGGCGCCCATTTA 8395
      |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 849 GAAAGGTGACAAGCATGCGGAGGTTCAAGCACTTGATATGCAACAAANAATGCTGANG 908
      |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 8396 AATATGTGAAGCTCATGCGAAGTTCATGCTCATATGCTGAGACATGACAGAG 8455
      |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 909 GTGCGACATTTATATTAAGTTAGACCATGTAGTATTTGGTTCAACACCACTGTG 968
      |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 8456 GTGCCGACATTTACGTTACACTCGAACCGTGCAGCATTAACGAAACACCGCATGTG 8515
      |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 969 TTACAAATTAATGATGTAGATGACANAAGT 1002
      |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 8516 CAGAAATTGATTATCAACTCTGATATCAAAAGAGT 8549
      |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

RESULT 8
BSUB0013/C BSUB0013 218470 bp DNA BCT 26-NOV-1997
DEFINITION Bacillus subtilis complete genome (section 13 of 21): from 2395261
to 2613730.
ACCESSION Z99116 AL009126
VERSION Z99116.1 GI:2634723

KEYWORDS

SOURCE

ORGANISM

REFERENCE

Bacillus subtilis.
Bacteria, Firmicutes; Bacillus/Clostridium group;
Bacillus/staphylococcus group; Bacillus.
1 (bases 1 to 218470)

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G.,
Azevedo, V., Bertero, M.G., Bessières, P., Bolotin, A., Borchert, S.,
Boris, R., Boursier, L., Brans, A., Braun, M., Brinell, S.C.,
Bron, S., Brouillet, S., Brusch, C.V., Caldwell, B., Capuano, V.,
Carter, N.M., Choi, S.K., Codani, J.J., Connerton, I.F., Cummings, N.J.,
Daniel, R.A., Denizot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D.,
Emmerson, P.T., Entian, K.D., Errington, J., Fabret, C., Ferrari, E.,
Foulger, D., Fritz, C., Fujita, M., Fujita, S., Fuma, S., Galizzi, A.,
Galleron, N., Ghim, S.Y., Glaser, P., Goffeau, A., Golightly, E.J.,
Grand, G., Giuseppe, G., Guy, B.J., Haga, K., Halech, J., Harwood, C.R.,
Henaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M.F.,
Ihaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y.,
Klaer, Blanchard, M., Klein, C., Kobayashi, Y., Koetter, P.,
Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A.,
Lardinois, S., Lauber, J., Lazarevic, V., Lee, S.M., Levine, A., Liu, H.,
Masuda, S., Mauel, C., Médigue, C., Medina, N., Meliardo, R.P.,
Mizuno, M., Moestl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M.,
Ogawa, K., Ogiwara, A., Oudega, B., Park, S.H., Pario, V., Pohl, T.M.,
Potterelle, D., Potvollik, S., Prescott, A.M., Prescan, E., Pujic, P.,
Punelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M.,
Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadale, Y., Sato, T.,
Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J.,
Sekowska, A., Seror, S.J., Serror, P., Shin, B.S., Soldo, B.,
Sotokin, A., Tacconi, E., Takagi, T., Takahashi, H., Takemaru, K.,
Takeuchi, M., Takemoshi, A., Tanaka, T., Terpetra, P., Tognoni, A.,
Tosato, V., Uchiyama, S., Vandenberg, M., Vannier, F., Vassarotti, A.,
Viari, A., Wambutt, R., Wedler, E., Wedler, K., Weitzenecker, T.,
Winters, P., Wiputt, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K.,
Yoshida, K., Yoshikawa, H.F., Zumschein, E., Yoshikawa, H. and
Danchin, A.

TITLE

The complete genome sequence of the gram-positive bacterium
Bacillus subtilis

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 218470)
Kunst, F., Ogasawara, N., Yoshikawa, H. and Danchin, A.

Submitted (18-NOV-1997) I. Moszer, A. Danchin, Institut Pasteur,
Regulation de l'Expression Genetique, 28 rue du Docteur Roux, 75724
Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr,
adanchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45
68 89 48

FEATURES

source

Location/Qualifiers
1. .218470
/organism="Bacillus subtilis"

terminator

/strain="168"
/db_xref="taxon:1423"
complement(7. .18)
/gene="cmk"

gene

complement(7. .696)
/gene="cmk"

CDS

complement(22. .696)
/gene="cmk"

/EC-number="2.7.4.14"
/note="alternate gene name: jofC, ypfC"
/codon_start=1

/transl_table=11
/product="cylidylate kinase"

/protein_id="CAB14221.1"
/db_xref="GI:2634724"

/db_xref="SWISS-PROT:P38493"
/translation="MEKKLSIAIDGPAAGSTVAKTIAEKKSYIYIDTGAMRYATY

AAIDENYDDEKLAELKRTDLEITTKGCVVFNCGDVPFAIRPDESNVSA
AKHSVREEMKROQUGEGGVWGDIDITHVDPNKAIFLLAIVEEPAKRIZE
NVKRGFVNYETLIEIARNDKDSERVSFLRAEDALEIDTTSLSIOEVADKILEA
VEOKSR"

gene
complement(775. .951)

/gene="ypfB"
complement(775. .951)
/gene="ypfB"

/function="unknown"
/note="alternate gene name: jofB"
/codon_start=1

/transl_table=11
/protein_id="CAB14222.1"

/db_xref="GI:2634725"
/db_xref="SWISS-PROT:P38492"

/translation="MKEFERLLIKLFIOAIIILGVQFLFHYQHIEPYISKYIQYEGV
DKKEENNRLETKH"

complement(96. .1649)
/gene="ypfA"

complement(96. .1649)
/gene="ypfA"

/function="unknown"
/note="alternate gene name: jofA"

/codon_start=1
/transl_table=11

/protein_id="CAB14223.1"
/db_xref="GI:2634726"

/db_xref="SWISS-PROT:P38491"
/translation="MIEGENVLEIIEENELKAKKAKVSTIENNELLAIVYDVYTG

RIVILNDMEYVEYFKDEVPYRISRIKGVKDKLQMICLEMPREKRIQROY
VRIDAVLDVQIOFGNEEIRTLISNAGSIVLADGSLFQSGESLRILRPEREH
TRILEAVAVRRIFNDPKSEKRMLETSEIAAGDQALLQYCIROLNRRKRAWE"

complement(1711. .3094)
/gene="ypbB"

complement(1711. .1733)
/gene="ypbB"

complement(1742. .3094)
/gene="ypbB"

/function="unknown"
/note="alternate gene name: jofB, yzua

similar to hypothetical proteins
/codon_start=1

/transl_table=11
/protein_id="CAB14224.1"

/db_xref="GI:2634727"
/db_xref="SWISS-PROT:P38490"

/translation="MIRGLIAVLGAIVGTGYGKKEOEKDAVLLHAENNYORAFH
ELTYVQDLHDKIGITLANNOSKSLSPALIDWRITSEANVSQPLTLPMPNTEE
LLSKIDFQSKTSRDLDOKPLDKNETSLINKLYQOSEDIONELRHVQLVMSKNLRM
MDVEMALASDEKSDNTIINSFTEKNGAFSTGDLSPFTSTKEKSGSHLKG
QISEQAKOIAERFAPDNDVSTIKVYSGKTRNDVYSIMKDPDKAVIYMDITKKG
HPYVLQNEVYKQKISLNDGSRALAFLLKNGEFDDLEIDESAGYDKIGVSVYVP
ENKVRMYPFAIRKRVALDGEVYGFARDFLTSKRRTLPKPAITEAKSKLUNKVO
VRETRLALITNELGQEVLCYEMGLTENIDTFMYINAKGSEKVEKLNABEYKDL

complement(3129. .4046)
/gene="sleB"

complement(3129. .4046)
/gene="sleB"

/note="alternate gene name: ypeA"
/codon_start=1

/transl_table=11
/product="spore cortex-lytic enzyme"

/protein_id="CAB14225.1"
/db_xref="GI:2634728"

/db_xref="SWISS-PROT:P50739"
/translation="MKSKGSIMACILIFFTITFTFINTETISAFNSNOVYORGATGDV

VELQARLQNGVYNGKIDGVGMYGAWAIVRQDQGLKEVGLVAGAKTQILLCKSK
YREYVMDOLNKNTFTTHYKIDPKYOTKPSKAOAKROQAEARQPAEFTCKPK
ANAKQONNTPKAKRODAVAAAMPQGFSSNDIRLQAQVYGEAREGPELGOVAITAY
ILNRNSPLFNPNSVAGVIEPLAFTVAGQITMOPNETAREAVLDAINGWDPSEBAL
YYPNPATSPWIGRPOIKRIKHLFCE"

complement(4162. .4841)
/gene="ypdC"

complement(4162. .4841)
/gene="ypdC"

complement(4162. .4841)
/gene="ypdC"

complement(4185. .4841)
/gene="ypdC"

complement(4185. .4841)
/gene="ypdC"

complement(4185. .4841)
/gene="ypdC"

complement(4185. .4841)
/gene="ypdC"

complement(4185. .4841)
/gene="ypdC"

```

/function="unknown"
/Note="similar to hypothetical proteins"
/codon_start=1
/transl_table=11
/protein_id="CAB14226.1"
/db_xref="GI:2634729"
/db_xref="SWISS-PROT:P50738"
/translation="MEATISAGIAPGIALISYVILKDOYDNEPVMVLRSPFLGVLV
PIMFIOYVLEKENYGGSEFSPSSGLESLSKFMIMISYVHAHDEYDITV
CASVSGFRATLENITLYLGHGVEHAFVRLLEVSCHALIGVMGTGKARSADAKR
VKWLTLSLVPSLHGSYDFILTALSNMYYMLPMVFLMWGLRKKAKARSVNMV"
gene
complement(4929..5935)
/gene="ypda"
complement(4929..4953)
/gene="ypda"
complement(4961..5935)
/gene="ypda"
/function="unknown"
/Note="similar to thioredoxin reductase"
/codon_start=1
/transl_table=11
/protein_id="CAB14227.1"
/db_xref="GI:2634730"
/db_xref="SWISS-PROT:P50736"
/translation="MIOGKATIIIGGPGCLSAIHLKOIGDALVIEKNVNSIYNY
PWHOFESSSEKLEIGDVAFTENKRPRIQALSYREVRKKNTRVNAFEMVRVTK
TONRTFVETSKETITTPYCIATIGTIDPNTMGVPEGLVFEHFEKGGHYFKDY
VVIKGNSSVDALVELVKSARVTVLVRKNEISPKPWLEFALVRNGTIRNEFG
ACVEKITENEVVERSEKELITIKNDPFAVMGYPHDQFLKIGVEIDKDEGRPFEN
EETMETNVEGVFIAGVIAAGNANNEIFENGFRHGHGIIAEIARENH"
terminator
complement(5995..6016)
/gene="ypca"
complement(5995..7324)
/gene="ypca"
complement(6044..7324)
/gene="ypca"
/function="unknown"
/Note="similar to glutamate dehydrogenase"
/codon_start=1
/transl_table=11
/protein_id="CAB14228.1"
/db_xref="GI:2634731"
/db_xref="SWISS-PROT:P50735"
/translation="MAADRNCHTEDEKIDVLKSTQVYIHRKLEKLGPEEYELKE
FMRLLTVIPVPMDDGSVKIFGTGRHNDVGPYTGKIREHNVTEKEKAKVALSIW
MSLKCGIIDLPGGKGGIVCDPDMSPRELERLSRGVRAISOVGTCKVPADVF
TNSOIMAMMDEYSRIDEFNSPGFITGRPLVVGSHGRESATKAGVTICIKEAARKR
IDIKGARVVQGFNAGSYLAKFMHDAGKVVGISDAYGLDPEGLDIDYLLDRDS"
Query Match      8.1%; Score 103; DB 2; Length 218470;
Best Local Similarity 60.6%; Pred. No. 4.6e-10;
Matches 166; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

```

```

RESULT 9
BARIBGENS
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
JOURNAL
COMMENT
FEATURES
source
promoter
gene
CDS
gene
CDS
gene
CDS

```

BARIBGENS 4286 bp DNA BCT 07-AUG-1997
B.amyloliquefaciens ribB, ribG, ribA, ribH & ribT genes.
X95955
X95955.1 GI:1592687
deaminase; GTP cyclohydrolase II; Iumazine synthase; reductase;
ribA gene; ribB gene; ribG gene; ribH gene; ribD gene; ribolavin biosynthesis
operon; Riboflavin synthase; ribT gene.
Bacillus amyloliquefaciens.
Bacillus amyloliquefaciens
Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillaceae;
Bacillus.
1 (bases 1 to 4286)
Gusarov,I.I., Krenova,R.A., Podchernyaev,D.A., Iomantas,Y.V.,
Abalakiina,E.G., Stoinova,N.V., Perumov,D.A. and Kozlov,Y.I.
Riboflavin biosynthesis genes of Bacillus amyloliquefaciens:
Primary structure, arrangement and regulation
Mol. Biol. 31, 446-453 (1997)
2 (bases 1 to 4286)
Gusarov,I.I.
Direct Submission
Submitted (27-FEB-1996) I. Gusarov, GNigenetika, 1st Dorozhny pr.
1A, 113545, Moscow, RUSSIA
On Oct 4, 1996 this sequence version replaced gi:1212772.
Overlapping sequence 227438.
Location/Qualifiers
1..4286
/organism="Bacillus amyloliquefaciens"
/strain="A 50"
/db_xref="taxon:1390"
321..348
/note="Riboflavin operon"
653..1768
/gene="ribG"
653..1768
/gene="ribG"
/codon_start=1
/transl_table=11
/product="deaminase"
/protein_id="CA65189.1"
/db_xref="GI:1592688"
/db_xref="SWISS-PROT:P70814"
/translation="MEETYNATLELARSGCGTQPNPLGAVVYKKRQIVGCAHLD
YGEHAHEVHAJNMGSIAKAGADLYVTEPCSHYKTPPCAEILMKSGIKRVLVEDP
NPLVAGKGIIMLEAGIEVTKGLRQOAEELKMFHMRGLPYVTLKAAASLDGKT
ATERGDSKWIITSEARLDQOYRSHORILVAGATVADNPSLCRLPDAYAKOVRVI
LDRLVPEETANVLTGDAAPMTWFTAGSDVRRKDRLLTAIGIKVETLETDRIHPEVL
SLIANGMSYVYVGASVHSGFYKACGFDLHFRFAPILIGTILAPSLISGEGFQSM
KDPHLOFTQITQIGPDIKLTAFKRDGDDVYRNR
1749..2396
/gene="ribB"
1749..2396
/gene="ribB"
/codon_start=1
/transl_table=11
/product="Riboflavin synthase (a-subunit)"
/protein_id="CA65190.1"
/db_xref="GI:1212774"
/db_xref="SWISS-PROT:O44680"
/translation="MFGIVVECTQIAIKTGLSMALTTAASKVTSQVRLGDSIAVN
GICITVGFSDNOTVDVMPETVYATSLNSKSKSKYLNLEAMKANGRFGRFVSGHY
DGLAEIRIEKKSNAVYIDLKLSPELTKTLVKSITVDGVSITIFGSDSESVSYA
PHIISETIFRKAAGSIVNIECDMIGRYLRFLLKRTBOTKSNQITTEAFSENGF"
2411..3607
/gene="ribA"
2411..3607
/gene="ribA"
/note="GTP cyclohydrolase II"
/codon_start=1
/transl_table=11


```
|||||
Db      9385 ATCATATGTGGCTATGATGACATCTACAGCCGACAG-TATAGCTGATGGAG 9327
QY      585 AA 586
Db      9326 AA 9325

RESULT 11
AE000675/c 14364 bp DNA BCT 25-MAR-1998
LOCUS      Aquifex aeolicus section 7 of 109 of the complete genome.
DEFINITION AE000675 AE000657
ACCESSION  AE000675.1 GI:2982863
VERSION
KEYWORDS
SOURCE
ORGANISM   Aquifex aeolicus.
            Bacteria; Aquificales; Aquificaceae; Aquifex.
REFERENCE
AUTHORS    Decker,G., Warren,P.V., Gaasterland,T., Young,W.G., Lenox,A.L.,
            Graham,D.E., Overbeek,R., Snead,M.A., Keller,M., AuJay,M.,
            Huber,R., Feldman,R.A., Short,J.M., Olson,G.J. and Swanson,R.V.
            The complete genome of the hyperthermophilic bacterium Aquifex
            aeolicus
            Nature 392 (6674), 353-358 (1998)
TITLE
JOURNAL    Nature 392 (6674), 353-358 (1998)
MEDLINE
REFERENCE  98196666
AUTHORS    2 (bases 1 to 14364)
            Decker,G., Warren,P.V., Gaasterland,T., Young,W.G., Lenox,A.L.,
            Graham,D.E., Overbeek,R., Snead,M.A., Keller,M., AuJay,M.,
            Huber,R., Feldman,R.A., Short,J.M., Olson,G.J. and Swanson,R.V.
            Direct Submission
            Submitted (25-JUL-1997) Diversa Corporation, Genomics, San Diego,
            CA 92131
COMMENT
FEATURES
SOURCE
            1..14364
            /organism="Aquifex aeolicus"
            /strain="vfs"
            /db_xref="taxon:63363"
            72..410
            /gene="rpsp"
            /note="ag_123"
            72..410
            /gene="rpsp"
            /codon_start=1
            /transl_table=11
            /product="ribosomal protein S16"
            /protein_id="AAC06492.1"
            /db_xref="GI:2982872"
            /translation="MAVRIRLAKFERKHHPIRIVYMDAKSPREKQYIDICTYPPKR
            KVLINVPKYEKAVLKGVELSHRAKAILMNGILKEVVPGEYEMKRVGDYVFEKRE
            SKSKSGGEAA"
            407..646
            /gene="ag_124a"
            407..646
            /gene="ag_124a"
            /codon_start=1
            /transl_table=11
            /product="hypothetical protein"
            /protein_id="AAC06503.1"
            /db_xref="GI:2982883"
            /translation="MSALKDIYELTAKELVDNKKVRYVTEIGEKTVVIELRVPDEL
            GKVIGKGRIRIAALRTILTAIGRKIGKRVLEIL"
            646..1044
            /gene="ag_125"
            646..1044
            /gene="ag_125"
            /codon_start=1
            /transl_table=11
            /product="putative protein"
            /protein_id="AAC06497.1"
            /db_xref="GI:2982877"

            /translation="MKERLALALLSFFLGSGSFSEFYVAPTLFLKLEBROQAGAVEKY
            FPYVEGIIILVGLSLFLGRDGRGLSLGILNLLLLLOEIVIPKJLHGKATNYEL
            FLKHGVSGMTINLAILLLTLGKVLILIKR"
            1049..1492
            /gene="ag_126"
            1049..1492
            /gene="ag_126"
            /codon_start=1
            /transl_table=11
            /product="putative protein"
            /protein_id="AAC06498.1"
            /db_xref="GI:2982878"
            /translation="MNKEGVKELISYGNMLAELEPINTKLEIKLKAQIIKTL
            LDVQKTEKEMKMDLEGVNEVAIEAKEMTONRMEYRMFLPEESSEVLKERMEG
            DIEDKHELSOTLGKISKLEKRLDLKDKGKILKIRRLIFR"
            1492..2310
            /gene="ag_127"
            1492..2310
            /gene="ag_127"
            /codon_start=1
            /transl_table=11
            /product="putative protein"
            /protein_id="AAC06499.1"
            /db_xref="GI:2982879"
            /translation="MLSELYVYLAMAFQAFPGAGIFIVIGILSLEPEKTVIYSS
            VVNLIGIISMLFIKTIIRPNKILFPLIAGSVAGIGVSAKLMEIDREVLKALIGD
            VLFLGITYDFLYQTSNLSKRLKASPPVCGFTGFLSGIFGILGMGCPPPVYLVNOYCRD
            LNTKTLISFFFAINVFRILFYLLGSTEWSLELIPALGLGALGVFAIYSRYF
            TPYFKRPLISVLSLGLFLEIGKEPSHDFHSHVIEESPYGGLSKSHKTPGN"
            complement(2134..3552)
            /gene="ag_128"
            complement(2134..3552)
            /gene="ag_128"
            /codon_start=1
            /transl_table=11
            /product="hypothetical protein"
            /protein_id="AAC06502.1"
            /db_xref="GI:2982882"
            /translation="MQRITVNPNEPYLSVIRKVKLSPIIVVNLITVENMISNLY
            SSISPSAAVATGFSLSLMTFYSLSLAUSYSGTINLIQAFVAKKDPSTILNGJLSR
            LISLPLEFYGRDFVLEFMKVLGASETVASLKEVLTPLFMPPIGILNTPTGYAVNGA
            GDRTKPMKVALIMNLTNIGTAVTTLNGVFGPKKGVEGAGCIIASELAEFIYTFLL
            IEFKPEPSHLRLERKILFKMKVRGLGTALRRATITTSFNVEVGLAKFGDKVLAHQ
            IGRRESPEMIGFGVMASTTLAGONYGARNYGMVAVTSAHFTLVNLSLGLIT
            ILPHRYLVPPSRDPEVLEMASTYLIQIGISQPMANASTISGALKMGKTHIRLFPVN
            ISSFWLERIIPSYFLKLVHSPLVPMGFMTEFVRALEFYTFKRVGKLLQTFQKE
            EEPGEDAEGDKPLDEVRSKYT"
            3567..4031
            /gene="ribh"
            3567..4031
            /note="ag_132"
            /gene="ribh"
            /codon_start=1
            /transl_table=11
            /product="riboflavin synthase beta subunit"
            /protein_id="AAC06489.1"
            /db_xref="GI:2982869"
            /translation="MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDCIVRGR
            EEDITLVAVPGSWEIPVAVAGELARKEDIDAVIAIGVLRGATPHFDYIASVSGLAN
            LSELRKPIITFGVITADTLFOAIRAGTGHGKMGEMALSAIEMANLFKSLR"
            4028..4474
            /gene="nusb"
            /note="ag_133"
            4028..4474
            /gene="nusb"
            /codon_start=1
            /transl_table=11
            /product="transcription termination Nusb"
            /protein_id="AAC06491.1"
            /db_xref="GI:2982871"
            /translation="MRYRKGAARDPAFLVLYRMDLRGENPGELFKEVVEEKNIKNKDAY
            EYAKIVDTAVRHIEIDSIIEKHLGYSIDRLGTVENRNALRGVAELIIFLKSPEGR
            VFIDIVDKKVADEKAGKPFVNGVLSAIYKAYITSSKEKSLKSE"

```

gene .5131 4676. .5131
/gene="aq_134"
4676. .5131
/gene="aq_134"
/codon_start=1
/transl_table=1
/product="putative protein"
/protein_id="AAC06500.1"
/db_xref="GI:2982880"
/translation="MROYSVIGSSKASFEFEYFVRGKELAKNLIYVGGGRGME
AVCCGAEGLGITIGMPSIDGHEANPVDIKVGTGMNPNPLIVASSEKVLAVGSN
YGLISELAYALLICKVYGVYKTRHVEGEVDESLEMLKRVDFEGRLL"
5128. .6075
/gene="nuem"
/note="aq_135"
5128. .6075
/gene="nuem"
/codon_start=1
/transl_table=1
/product="NADH dehydrogenase (ubiquinone)"
/protein_id="AAC06490.1"
/db_xref="GI:2982870"
/translation="MKVFITGATGFCVGRHLYRELLNNGEYHAGVRLSKIERLFGNO
VKGIYVDEKDSIREALGVNPDVFIHLIGILTEKKKGITFERVYGHKNLVEVS
KGFNKKFLFMSALGTHDEAFSRHQTRMAERMYINSGLVYIIFRSIILIGPQKLF
EDMKIKRYLIPVVALPDGPNFQFQVPRVDACAYAAALKPEPDRKIYELCGKRVYT
FKELADIFSYMDRKKVLMIPMKRLMEFAGIIVKIIIEPPFSDDQIKMMKPNVCGV
LKDAISNGTICKRREPDIPEESLKNSLENFDIILIKG"
6098. .7165
/gene="cpv"
/note="aq_136"
6098. .7165
/gene="cpv"
/codon_start=1
/transl_table=1
/product="cytochrome c peroxidase"
/protein_id="AAC06485.1"
/db_xref="GI:2982865"
/translation="MRKGLIMAFGVYAKERIDDKELLMARQYFKPLPKYAENDP
ONPTEPKVKIGKMLYDPRLSKGLISCNCHMLRAGVNLSTISGHRALIPRANA
PTYNAALHINQFMDGAKRKYDEBOALGPYVPIEMANTEAVATVYVYVLFKK
AFPEKDDVKTENIGKAIGAFERTLMPSRDEFKGNTRALTQERKGLTPIEVGC
VACINPGVQGNMFKAQGMITEYKVTYPIYLVGKPAIKDGFQVTKEDMFVKR
VPSLRNIETHYPIFYHDSVMSLEDVAVRIMAKTQLGKELTDQVNDIVAFKALTKIP
KHALVEPELPSTKTPPEL"
complement(7162. .8222)
/gene="ribD1"
/note="aq_138"
complement(7162. .8232)
/gene="ribD1"
/codon_start=1
/transl_table=1
/product="riboflavin specific deaminase"
/protein_id="AAC06487.1"
/db_xref="GI:2982867"
/translation="MIREIDKNYKMLAISLAKRRKGYTHPNPTVAGVVEKGIYGLG
YHEAGKPAHEVMALQAGEKAKGATLYTLEPCTHFGRPCTDAIRSGIKKVVAA
TLDNPLMSGKGVKLRNAGIEVDVGEEDARLNEDEFYITQERPYITLKAAOTLI
DGKATLTGSSKMTKTSKSRKVAHLIRREATAVLAVGNTVTKDDPHLIVRPEKDP
VRITLDPELYPLSAKULNTEARPTVITTKENKAEKLEKQYVILIKFNKNIILII
KRLEULETLMHLMVEGGPRTLTSLKEGFDRIVVFIAKILWEGSLUSDGINSIEES
LKRVKKRYENLGEDVLVFFKRY"
complement(8229. .9173)
/gene="ribP"
6.7%: Score 84.8; DB 1; Length 14364;
Best Local Similarity 49.0%; Pred. No. 1.6e-06;
Matches 218; Conservative 0; Mismatches 227; Indels 0; Gaps 0;

QY	789	CACCGTTGGGGCGTGTATAGTTATGAAGGATAGATTGTTGGTATGTGTCGACACTTGA	848
Db	8150	CTACGGTGGGAGCGGTTTGTCTAAAGGAAGGAAAAATCGTGGGATTAGTATATCGAGAA	8091
QY	849	GAAGAAGTGAACAGATCGGAGGTTTCAAGCACTTGATGTGTCACACAMNATGCTGAAG	908
Db	8090	AAGCAGAAACCCCTACCGCTAGATTATGGCTCTCGTCAGCGCGGGGAAAAAGCAAAAG	8031
QY	909	GTGCGACGATTTATATACGTTAGACCATAGTCACTTTTGGTTCAACACCACCTGTG	968
Db	8030	GGGCAACCCCTTACGTTACCTTGAGCCGCTGACTACCTTCGGGAAACCTCCTCCGTCA	7971
QY	969	TTACCAAAATATATGATTGTAAAGTAAAGACNAGTGTATTCNCAACNAAAGACATTC	1028
Db	7970	CGAGCGCATATAAAGAGCGGATTAAGAGGGTTGTAAGTGCAACGCTTGACCCAAATC	7911
QY	1029	GTTACACACACATGGGTATGATAGACGTACCGGGCGTCCACGGTATTTAGGGTTGAATG	1088
Db	7910	CTTTGATGTTCTGGAAAGAGTGTGAAAAACGTAGAAATCGCGGATTAGAAGTGAATGTC	7851
QY	1089	CGTTGATGATGAACGGGCATCACATTATACCAAAAGCTTTTTTAAAGCAAAAGCAA	1148
Db	7850	GTTGTGCGAAGAGGAGCGCCAGAGCACTCAAGAGATTTTTTTTCTTACATCAACTCAG	7791
QY	1149	AGCAACTTGCACCAAAATTTACAGTG	1173
Db	7790	AAAGACTTACATACGCTAAATG	7766
RESULT	12		
LOCUS	U32775	18073 bp	DNA
DEFINITION	Haemophilus influenzae Rd section 90 of 163 of the complete genome		
ACCESSION	U32775	142023	
VERSION	U32775.1	GI:1573951	
KEYWORDS			
SOURCE	Haemophilus influenzae Rd.		
ORGANISM	Haemophilus influenzae Rd Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; Haemophilus.		
REFERENCE	1 (bases 1 to 18073)		
AUTHORS	Fleischmann,R.D., Adams,M.D., White,O., Clayton,R.A., Kirkness,E.F., Kerlavage,A.R., Bult,C.J., Tomb,J., Dougherty,B.A., Merrick,J.M., McKenney,K., Sutton,G.G., FitzHugh,W., Fields,C.A., Gocayne,J.D., Scott,J.D., Shiley,R., Hu,B.I., Glodex,A., Kelley,J.M., Weidman,J.F., Phillips,C.A., Spriggs,T., Hedblom,E., Cotton,M.C., Utterback,T., Hanna,M.C., Nguyen,D.T., Saudek,D.M., Brandon,R.C., Fine,L.D., Fritchman,J.L., Fuhmann,J.L., Geoghagen,N.S., Gnehm,C.L., McDonald,L.A., Small,K.V., Fraser,C.M., Smith,H.O. and Venter,J.C.		
TITLE	Whole-genome random sequencing and assembly of Haemophilus influenzae Rd		
JOURNAL	Science	269 (5223), 496-512	(1995)
MEDLINE	9530630		
REFERENCE	2 (bases 1 to 18073)		
AUTHORS	Tatusov,R.L., Mshnegian,A.R., Bork,P., Brown,N.P., Hayes,W.S., Borodovsky,M., Rudd,K.E. and Koonin,E.V.		
TITLE	Metabolism and evolution of Haemophilus influenzae deduced from a whole-genome comparison with Escherichia coli		
JOURNAL	Curr. Biol.	6 (3), 279-291	(1996)
MEDLINE	96398784		
REFERENCE	3 (bases 1 to 18073)		
AUTHORS	White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-JUL-1995)	The Institute for Genomic Research,	9712
	Medical Center Dr, Rockville, MD 20850, USA		
REFERENCE	4 (bases 1 to 18073)		
AUTHORS	White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-SEP-1997)	The Institute for Genomic Research,	9712
	Medical Center Dr, Rockville, MD 20850, USA		
REMARK	The H. influenzae sequence has been updated by R. Fleischmann. New database matches have been assigned, product names have been		

Improved, and a number of frame shifts have been corrected. We gratefully acknowledge the work of Tatusov et. al. We have incorporated their annotation into the /notes fields of the corresponding H. influenzae genes
(bases 1 to 18073)

AUTHORS White, O., Clayton, R.A., Kerlavage, A.R., Fleischmann, R.D., Peterson, J., Hickey, E., Dodson, R. and Gwinn, M.

TITLE Submitted (28-MAY-1998) The Institute for Genomic Research, 9712

JOURNAL Medical Center Dr, Rockville, MD 20850, USA

REMARK The whole genome was shifted by 588 nucleotides for a new start
On Sep 30, 1996 this sequence version replaced gi:1221677.

FEATURES
location/Qualifiers

1..18073

/organism="Haemophilus influenzae Rd"

/db_xref="taxon:71421"

/complement(146..1327)

/gene="HI0929"

/complement(146..1327)

/note="similar to GB:M77129 SP:P24196 PID:146678
PID:882568 GB:U00096 percent identity: 41.10; identified
by sequence similarity; putative"

/product="conserved hypothetical protein"

/transl_table=11

/protein_id="AAC22589.1"

/db_xref="GI:1573952"

/translation="MKRVTEFPTRPDMVQQLLVGVEDYVLPSSDGSVHWSNVAYEE
TLAIDRIEDTNEHLMSCDLPADAEIKKGDYENHETELKQDIEISMNOPIYLG
RFDFYGDNDLKMFEYNADPTSLLEAVVOMOLEQIEGLKRDQFNHHELIKFE
QFLKQSGKTDPLSAMQDAGREDGWDVLDVAVYAGNINHQLAIVEDIGYSETK
FVLDNQPIFMLEFKLYPLEVLSAEPRARHTTAPRIEPAKMLSKKLKLKLNAR
YFNHPLPAFTPEFLPKDLSMVKPPLGIRGANVTEKNGVEFAKGSSEHST
YKNSGYVQOKFELPSFGDMYPIIGSWVGVDAWGMLREDFTAVTGNDSHFIPHYFV
P"

/complement(1328..1948)

/gene="HI0930"

/complement(1328..1948)

/gene="HI0930"

/note="hypothetical protein; identified by Genemark;
putative"

/codon_start=1

/transl_table=11

/product="H. influenzae predicted coding region HI0930"

/protein_id="AAC22600.1"

/db_xref="GI:1573966"

/translation="MARRKNSIKIENNEKNOILVSLIALLGCGSEBOVORDVYO
SLDDCLADKKIKELCEADKNTESOKTATTOOGLGINRDNGNAESAVKPNENNAO
ANSENRAESTRTAESRPSISGAIAGGVGYMAAPRISFLGSPYHGRARVTTPTG
QVAVQPTNRSGVGMVLVKGAGSNKSPVSRGGSFSPNKRTHRSSGC"

/complement(1921..2406)

/gene="HI0931"

/complement(1921..2406)

/note="hypothetical protein; identified by Genemark;
putative"

/codon_start=1

/transl_table=11

/product="H. influenzae predicted coding region HI0931"

/protein_id="AAC22601.1"

/db_xref="GI:1573967"

/translation="MKSHTRLYKSAVFLYIFGKRMHLTLEDELFLGTPKOVGTHS
TVEHFVAFMEDDEGEGFYALDMROMAOPIDVMLHAYVNDSTSNHHEARKLEICDE
SGVVALLLINGYHAAVDFARLVGYNSNKPDPDLMSMTREITNKQABQWLGVKTI
R"

/complement(2424..3734)

/gene="HI0932"

/complement(2424..3734)

/note="similar to GB:M12843 SP:P08324 GB:X82400 PID:147479
PID:563868 percent identity: 85.02; identified by sequence
similarity; putative"

/codon_start=1
/transl_table=11
/product="enolase (eno)"
/protein_id="AAC22590.1"

/db_xref="GI:1573953"

/translation="MAKIVKIGREIISDRGNPTVEAEVHLXGFEVGLAABSGASTG
SRALRLRQDMSRFGLKGVLAFAVANNIEAOATVGDATNOAEDIDIMIDGTEN
KSNFGANILAAVLANAKAASGLPLAYATIAELNGTAGVSMPLPMNIIINGEHA
DNVDIOEFMIQDPVAKTLREALRIGAEVFNHAKYKKAKGSTAVGDEGGAPNLAS
MDALALCKENAEKAGYVYKGVTLAMDACASSEFYKKEKNGHEMKKEGTSFOETH
YLEELCKEPIVYSIEDGQDSEWGEFAVQTKVGLGVQVLDVGLFTNTKILKEGLEK
GIANSILIFENQIGSLTERLAIKKAKDAGYAVISHRSGEDATADLADLVAGTAGO
IKTGSMSRSDRIKYNQLIRIEALELRAGTAPAFPLKAVKQGA"

/complement(3855..5060)

/gene="HI0933"

/complement(3855..5060)

/gene="HI0933"

/note="similar to SP:P37631 PID:912464 GB:U00096
PID:2367234 percent identity: 58.35; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=11

/product="conserved hypothetical transmembrane protein"

/protein_id="AAC22591.1"

/db_xref="GI:1573954"

/translation="MSQYSENIIGAGAGLFCQAOLAKGKSVTFVDNGKIKRKIL
MSGGRCNFTNLEVPAPYALSONPHVKSAIARVTMDPISLVAEGITVYHREKGL
FCDBGEQIVEMLKSECDKYGKILRESEVQERIQNDEKRFVQVASTONQCKNL
IVATGLSMPLGATPFGQIAEQFIPVIPPASVLPVTETDLFTLALSGISLPV
TITALGSKSFYNQDLTFHNGISGPAVALSNWQPSVEIDLPHNVEEELNOAKO
SRRKQKTLIVLELPLKLVLELIEQIVODEVIANISVRYKLVDFIHHHEFTPG
TEGYRKAETLVGVDFKIVKISVTMESNQVSGYFIDEVLDVTGMLGYNFQWAMSSAY
ACALSSRSQ"

/complement(5057..6211)

/gene="HI0934"

/complement(5057..6211)

/note="similar to SP:P32711 GB:X72298 PID:396410
PID:404307 GB:U00096 percent identity: 50.41; identified
by sequence similarity; putative"

/codon_start=1

/transl_table=11

/product="cytochrome C-type biogenesis protein (nrfp)"

/protein_id="AAC22592.1"

/db_xref="GI:1573955"

/translation="MKSIDITFRGFIQVQKYFFATALLPAFSLETOAMVDYTFQ
NDDRTVAELASLRCPQONONLVESNPAYDRLRVKWDGSKNOQIDIMT
ARQGDVYKRPPEKMTALMLLPVALLLAVALLFNSNRKQFSKVAOQLENDI
ISLPTFGSGSPKQSGPSKSGKVSRTYVLTLLIAPATYFSLDRFSRQGE
QSMIEGHNQNVENDENKENVTEKONKLRTPDNAETMLQGEAYVONNEFDSALV
AYSNAEKISGSKNITLIGLAATLXYQAQOMQKVOLINELAKKREVSISLAT
IALENROYQAGNYLQOQLDSSGAAVDRASVIOKMKMLDFLORGEQNP"

/complement(6208..6738)

/gene="HI0935"

/complement(6208..6738)

/gene="HI0935"

/note="similar to SP:P33926 PID:405920 GB:U00096
PID:178523 percent identity: 42.60; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=11

/product="thiol:disulfide interchange protein (dsbe)"

/protein_id="AAC22593.1"

/db_xref="GI:1573956"

/translation="MKKKLIFTPTLFFVLLGVCILLIAGINQDPKRTASALLDKPVPEF
YCANLHPSQIVSPKEPKOPFLNLVWMSGCKSEHLLIETATLPIVGVDRDN
RONGIKMLKMGNPFLITIDSHGELGKLGVGABETVLVDENGVIYRRHSGILDK
TWQTVFLPKITLALXNK"

/complement(6738..8645)

/gene="HI0936"

/complement(6738..8645)

/note="similar to SP:P32710 GB:X72298 PID:581146 GB:U00096

PID:1790511 percent identity: 42.42; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="cytochrome C-type biogenesis"
/protein_id="AAC23594.1"
/db_xref="GI:1573957"
/translation="MLPELGFLLILATASAFIALVPOGLFKKPTLINAAMPLSY
LPTATLSTIGLASFADDETLLEVVAHNSOLPFEKVAATGHEGSMLELFS
LSMLAFAFRNRKDRFSSQSLSLGLICGFVFLFISNPGRIFFPABESRD
NPMLODVLIFHPLLVGVYGFVAFVEMSSALINSAQIARSMKMWLVSWFL
TIGVIGAMWAVYELGMGWMFMDVENASLMPMLGLALHSLMATEKQVFSYWT

Query Match 5.5% Score 69.4; EB 65; Length 18073;
Best Local Similarity 49.7%; Pred No. 0.0011;
Matches 172; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

QY 657 ATAAATTTTCTCCTTGCATCTTAATCATGATGAGGATTTTCTTATAGAGTG 716
DB 15823 ATTGATCAATTACTATCGAAATTAAGAAATTAATGATGATTTTCT 15882
QY 717 ATCAATTGATCAATTATGATGATGATGATGATGATGATGATGATGAT 776
DB 15883 CACAAGATTGCGTATTATGACACGCTTAGATTGGCAGCCAAAGGTCAATATACCA 15942
QY 777 CAGGTGTTAATCCACCGCTGGCGTGTAGTTATGATGATGATGATGATGATG 836
DB 15943 CTACACCGAACCATTGGTGGGATGCGTATGTTAAATGTGAATTTGGGGAGG 16002
QY 837 GTGCACACTTGAGAAAGGTGACACGATCGAGGTTCAAGCACTGATGACACAC 896
DB 16003 GCTTTCATTTTAAGCAGACGACCTGCTGACGCGTGTAGCTTACGCTGATG 16062
QY 897 AANAATCTAAGTGGCAGATTTATATAGTTAGCAGGATGATGATGATGATG 956
DB 16063 AAAATGCTAAGAGACACCGCTTACGCTACCTTACGCTTACGCTTACGCT 16122
QY 957 CACACACCTGCTTACCAAAATATGATGATGATGATGATGATGATGATGAT 1002
DB 16123 CACACCTGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCTTAC 16168

RESULT 13
AE004298/c DNA BCT 31-JUL-2000
LOCUS AE004298 AE003852
DEFINITION Vibrio cholerae chromosome I, section 206 of 251 of the complete
chromosome.
ACCESSION AE004298.1 GI:9656821
VERSION
KEYWORDS
SOURCE
ORGANISM
Vibrio cholerae.
Vibrio cholerae
Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

REFERENCE 1 (bases 1 to 11695)
AUTHORS Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwin, M.L.,
Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Unayam, L.A.,
Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,
Ernst, A.E., Vamathevan, J., Bass, S., Qin, H., Dragol, I.,
Sellers, P., McDonald, L., Uitterlinden, T., Fleischmann, R.D.,
Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R.,
Mekalanos, J.J., Venter, J.C. and Fraser, C.M.
DNA Sequence of both chromosomes of the cholera pathogen Vibrio
cholerae
Nature 406, 477-483 (2000)
2 (bases 1 to 11695)
AUTHORS Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwin, M.L.,
Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Unayam, L.A.,
Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,
Ernst, A.E., Vamathevan, J., Bass, S., Qin, H., Dragol, I.,
Sellers, P., McDonald, L., Uitterlinden, T., Fleischmann, R.D.,
Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R.,
Mekalanos, J.J., Venter, J.C. and Fraser, C.M.
Direct Submission

JOURNAL Submitted (14-JUN-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
source
1. 11695
location/Qualifiers
organism="Vibrio cholerae"
strain="N16961"
serotype="O1"
/db_xref="taxon:666"
/note="biotype: E1 Tor"
complement(97.825)
/gene="VC2260"
complement(97.825)
/note="similar to GB:D26562 SP:P07906 GB:M5106 PID:146727
PID:473823; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="ribosomal protein s2"
/protein_id="AAP95404.1"
/db_xref="GI:9656822"
/translation="MASVSRDMUTAGVHGHOTRYNPKMKOIFGARRVHINLE
KTPMFENDALAEALAKVEKKGVLPVCTKRAASVSVEALASQYVYNNRWGMIT
NMKVRSIRKLKELEVDGTFDKLTKRKALKRTEKLESLGIGIDKMGGLPDA
LEVIDADHEHAIKEANNLGI PVFAVVDITNSSPDGVYIIPGNDADIRAVQVLYNAAA
QAINEGRNKDVAAVAEKDCEVEAE"
1166.2008
/gene="VC2261"
1166.2008
/note="similar to GB:D26562 SP:P07906 GB:M5106 PID:146727
PID:473823; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="methionine aminopeptidase"
/protein_id="AAP95405.1"
/db_xref="GI:9656823"
/translation="MSIKIRNAVEIERMRYAGRLAEVLEIMEHRYKAVYTTLEDDI
CHKYITVOGAIAPALVNHGPKSICISINHYAGCPASEDEYFGIOPAVLRDD
LINDIVIKDYGDSKMEFLIGDIEGRKLCVHOECTYALAKOVKGVGLGIG
TTTEKHITKNKNRPKFSIVRDYCHGGAEPHERPOVYHKNRSRYLRBSMIFT
IEPIMNAGFEGRDDESMVTYADSKSAQWHTLLVATGCEIITLSEESLPRI
LNNA"
2098.4743
/gene="VC2262"
2098.4743
/note="similar to GB:M96431 SP:P27249 PID:146155 PID:49393
PID:473822; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="protein-P-II uridylyltransferase"
/protein_id="AAP95406.1"
/db_xref="GI:9656824"
/translation="MHRAMPYOSPTFOEPDLTVESLKOQLESFTEYOKOEFDDHP
VTVDLVRSEMDLILRLMWFEGFDELVEVLAAGVGSGELHPSIDIDLVSQ
PLSFOVANKISQFLTLMDLKLEIGHVAFYEOCAEIGKADLVATNLOEARLSCGE
EPPRIKRWIHSESPFSEIFYOAKVREOKRERHAYDNTNEPDIKSPGGGRDH
TLSVARRHGATSLYEMSRGFTLDAYRLEVSCOPILRYRVALRIELKRYNRT
FAHVOYARHGYGEBNKGLEMMKFEFRLRVALELKNKLITFKALINNGEAE
VAIIDDFORGNMIEARKPALFQARPEITLDMFLHNASSTIESVAPATMRQIAR
RLNKFPLHTLPAREKFLELVHNPALHKAFFSQNHKGLAAYLPQNOIVGQMPDL
FHVYTVDEHSIRLKHIFLSDANNNHRPCCCEIYPIKQKELLILAAIFHDGGR
GGSSEIGADAPDCEIEHGSKPEAKLVAMVYKNNKLVSTAGRDIDYDPDVIIEFA
KRVDERLLETIVADICATPELINSKRRLLLELYSTORALRGLENVDR
ERHNNQMASALLRKGEFSRELEVLMORKAYFLRHHTKQIAMWCTHLREDS
KPLVLSSKATRGCEVFIYTKDAALFAIVAAFLVAVLADPQAKVMAKSDGVLT
EMVLDNGOALIEDRHQALIRHLYHVEDLEPTQKRRIRPRNQHVKYQVDFLT
KSKRTIMEFVALDPLGLATVGFAPLNDLHAATITTGEEADELLITNAQGR
LNNEEDHLRKLLENHAEAPSQ"
complement(4818.4916)
/gene="VC2263"
complement(4818.4916)

gene
CDS

gene
CDS

/gene="VC2263"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAFG95407.1"
db_xref="GI:9656825"
/translation="MSLMGTHSLAAYLQLOLVWYFTFKALNLL"
4989..5402
/gene="VC2264"
4989..5402
/gene="VC2264"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAFG95408.1"
db_xref="GI:9656826"
/translation="MNVILIKRLPMFPHVLVGLIGINEPQIRSYSLRGAHKDKYLITF
KKKELSAKSVMKEFYPRQIKNLVDSSGHQYEVTETNRMLTLVIDLNKITPEWM
GEVDVKORILLDRLEKRVASAKIAIEADLOKLIN"
complement(5443..5952)
/gene="VC2265"
complement(5443..5952)
/gene="VC2265"
/note="similar to GB:U00096 PID:1773102 PID:1786620
PID:2329841; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="phosphatidylylglycerophosphate A"
/protein_id="AAF95409.1"
db_xref="GI:9656827"
/translation="MSDPRSMLSLSNPWHLATGFSGSGLPIVGTMGTLASVPFLTLM
LAQLWLVAVVVVFASLGICGVTSDDDVQHDSIVDEFAGFWITLIPLYQR
LPVPMKMVLVGFVLFRRFDMLKPMPISWLDMKRVRHGCGIMIDDLVAGVMAISIALMI
CYWLEMEDLV"

gene .5949 .6953)
CDS complement(5949..6953)
/gene="VC2266"
complement(5949..6953)
/note="similar to GB:U00096 SP:P77785 PID:1773101
PID:1786619 PID:2329840; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="thiamin-monophosphate kinase"
/protein_id="AAF95410.1"
db_xref="GI:9656828"
/translation="MQRFNQRYMMWFGEFNLDIXYSNRQAQRKDVIHLAQQCAIYK
VEPSNRVAISDTLVAEGHTHFLAQNAPPAWAARKALASNTSDLAAGATPPAWSVLTIPLP
EIDDAWLRPFCDAGEFAKYKNVOLIGDTRKGPSITLIYVQGLPREQAOLRSBGAJLTP
GDMLVTGGDISOAGLTVIIDPERKHLPFADIIDEORIHYLTSPRIYGVALVHAASA
IDSIGCLADLIQHIRSNNGCASIDVSLLPLSKTELQPVDSVTSAQQYALTGSBEYLL
CTFIPEERNGSEENLAHCGRKVTCIGQIRPAGTFELHNQSKLLMDQLAGTDHRKKVKA
"

gene
CDS

complement(7014..7484)
/gene="VC2267"
complement(7014..7484)
/gene="VC2267"
/note="similar to GB:M26839 SP:P04381 GB:X00681 GB:X64395
PID:146576; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="N utilization substance protein B"
/protein_id="AAF95411.1"
db_xref="GI:9656829"
/translation="MGASVKKPARARNARNOFALOAIYSWOITKENVATEIEDOLFSTGYX
DEEEHRAEPALAABEDVSYPFRDLACVYNHHDELSQLRPFSRPMODIMNELALL
LRLAYEMTRREDYVYKVVINAEALVELAKVFPAEDSHKFVNGLDKAAPHVKKKA"
complement(7484..8005)
/gene="VC2268"
complement(7484..8005)
/gene="VC2268"
complement(7484..8005)

/note="similar to GB:U11391 PID:496182 SP:PI1963; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="riboflavin synthase, beta subunit"
/protein_id="AA05412.1"
/db_xref="GI:9656830"
/translation="MESGSDSHLNRRKGLKMYIEGCFAPNKKIVISRFNSFIN
ESLISGAIIDLKRRGQISDNITVVRPGAVEPLVQARAKTDYDAIVSLGVIRGG
CFPFEDVCESEMNKGIAOVSLFEFIPAFGLITVDITDIOAIERAGKAGAEALSL
ALEMINVLEISDS"
complement(8219 . .9328)
/gene="VC2269"
complement(8219 . .9328)
/gene="VC2269"
/note="similar to GB:M90094 SP:Q02008 PID:150708;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="3,4-dihydroxy-2-butanone 4-phosphate
synthase/GTP cyclohydrolase II"
/protein_id="AA05413.1"
/db_xref="GI:9656831"
/translation="MPSIPEQIEEDIRCKMYILMDDEDRENGDILMAEHITPAA
INFNATHRGRLICLTITKRCRRGLGPMQDNNAAQYTNFTYSIEAEVGTGISAA

Query Match 5.4%; Score 68.4; DB 1; Length 11695;
Best Local Similarity 52.3%; Pred. No. 0.0017;
Matches 147; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

OY 727 TCAATTATGCGATTATGCGATTCAACTCCCAAAATATGCTACAAAGCTCANACAGGTGTTAA 786
||||| ||| | ||||| ||||| ||| | |||
Db 11104 TCAATGATGCTCGCGGATGAACTTCGTCGGCGGGGCTTTTACCAGTTCCTCTAA 11045
||||| ||| | ||||| ||||| ||| | |||
OY 787 TCCACCGCTTGGCGCTGTTAGTAAATGAGAGTATGTTGTTATGTCAGCACTT 846
||||| ||| | ||||| ||||| ||| | |||
Db 11104 TCTTATATGTCGGCTGCGATCCTCCTCGGCGTGAGCAGATTGAGGGAAGTTTCCATT 10985
||||| ||| | ||||| ||||| ||| | |||
OY 847 GAGAAAGGTGACAGCATGCGGAGGTTCAAGCATCTGATGCGACAAANATGCTGA 906
||||| ||| | ||||| ||||| ||| | |||
Db 10984 TCGCGCGGGCGCAACCCATGCTGAAGTGCATGCTATGCGCAAGCGGTGAGCTTACG 10925
||||| ||| | ||||| ||||| ||| | |||
OY 907 AGGTGCGAGATTATATATACGTTAGAGCCATGTATGTTGGTTCAACACCACTTG 966
||||| ||| | ||||| ||||| ||| | |||
Db 10924 CGGTGCGACCCCTTATGCACTTATAGACCTTGTCTCATTTATGTCGACACACCGCCTTG 10865
||||| ||| | ||||| ||||| ||| | |||
OY 967 TGTTACAAATATATGATTTAGATGATGCGANAGTACTAT 1007
||||| ||| | ||||| ||||| ||| | |||
Db 10864 TGCTGAAGGCTTATTAAAGCGGGGTTCGCAAGATGATT 10824
||||| ||| | ||||| ||||| ||| | |||

RESULT 14

LOCUS	ECNUB	2949 bp	DNA	BCF	26-JUL-1993
DEFINITION	E.coli genes nusB (ssyB) and 3 orfs.				
ACCESSION	X64395.S47077				
VERSION	X64395.1 GI:42147				
KEYWORDS	nusB (ssyB) gene.				
SOURCE	Escherichia coli.				
ORGANISM	Escherichia coli.				
REFERENCE	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.				
AUTHORS	Ito, K.				
TITLE	Direct Submission				
JOURNAL	Submitted (06-FEB-1992) K. Ito, Inst for Virus Research, Kyoto University, Sakyo-Ku, Kyoto 606-01, JAPAN				
REFERENCE	2 (bases 1 to 2949)				
AUTHORS	Taura, T., Ueguchi, C., Shiba, K. and Ito, K.				
TITLE	Insertional disruption of the nusB (ssyB) gene leads to cold-sensitive growth of Escherichia coli and suppression of the secY24 mutation				
JOURNAL	Mol. Genet. 234 (3), 429-432 (1992)				

	MEDLINE	930342316	location/qualifiers
FEATURES			1..2949
source			/organism="Escherichia coli"
			/strain="K12"
			/db_xref="taxon:562"
			/map="9.5 min"
			330..357
misc_feature			/note="promoter-like sequence p1"
			333..362
misc_feature			/note="promoter-like sequence p1"
			391..840
CDS			/note="orf1"
			/codon_start=1
			/transl_table=1
			/protein_id="CAA5734.1"
			/db_xref="GI:42148"
			/db_xref="SWISS-PROT:P25538"
			/translation="MHCPCFADVTKEIDSLVGESSVRRRCLVCNEPFTFEVAA ELVNPVYKSNDAVEPEPNEEKLRSGLMALEKRPYSDDVMALNHKLSQRAIGERD VPISMGNLVMEQLKDKKAVIRFASYSRSEDIEKERGETARLED"
			844..1947
CDS			/note="orf2, homologue to B.subtilis ribG"
			/codon_start=1
			/transl_table=1
			/protein_id="CAA5735.1"
			/db_xref="GI:581147"
			/db_xref="SWISS-PROT:P25539"
			/translation="MODEYYMARALKLAQRGRFTTHPNNGCVIVKDEIIVEGHQHQ RAGPHAEVHALRMAGEAKATAYVTLEPCSHHGPRPPCCDALIAGVARVASMODD RNIPVARGLYRLQOAGCIDVSHGLMSAEQDLNKGFLKRWGTGPYIOIKRGASLDR TNAAAGSORITTSPOARDVOLLRQAOSHAIITTSATYLADDPALTVKRSELDDETOL YPNOILROPRIIVIDSQNRVTPVHRIVQDPETVFARTQEDSRWMPETVRLTLPEKH GHLDVLVLMQLGKOQINSIWENAGPTLAGLLQAGLDELIVYIAERLLGSDBGIC TLPLEKRLADAPORFKFERIRHVGDPVLHIVA"
			1957..1986
misc_feature			/note="promoter-like sequence p2"
			2036..2306
CDS			/note="orf3, homologue to B.subtilis ribH"
			/codon_start=1
			/transl_table=1
			/protein_id="CAA5736.1"
			/db_xref="GI:42150"
			/db_xref="SWISS-PROT:P25540"
			/translation="MNITEANVAAPPDAVAITTFARNFNINDSLLEGALDALKRIGOV KDENTITVWVGAVEALEPLAAGLAKTKGYAVIALGVIRGSTAHFEFYVAGASNGILA HVADDESIPVAFGVLTTESTIEQALIERGSTAGNKGAEMALTALEMIVLKAIRA"
			2526..2945
gene			/gene="nusb (ssyb)"
			2526..2945
CDS			/gene="nusb (ssyb)"
			/codon_start=1
			/transl_table=1
			/protein_id="CAA5737.1"
			/db_xref="GI:581148"
			/db_xref="SWISS-PROT:P04381"
			/translation="MKPAARRARAECASVALLYSQWSQNDIADVEYOFLEADVDYD VLEPERELLAGVATNTAYLDLGIMKPYLSRLLEEIGOVERKAVIRLALIELSKSDVPYKVY AINNAIFLASFGAESDHSHKFVNGVLDAAPAIRNRKK"
			BASE COUNT 708 a 739 c 825 g 677 t
ORIGIN			
			Query Match 5.3% Score 67.6 DB 2 Length 2949;
			Best Local Similarity 52.3% Pred No.0.003;
			Matches 145; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
QY	729 AATTATGATTATCGCATTCACACTCCAAATATGTCACAAGGCACACAGGTGTTAATC	788	
DB	857 ATTACATGGCGCGGGCGCTTAAGCGTGCGGCCAACGACGAGCTTTTACCACGATCCCAACC	916	
OY	789 CACCCGTTGGCGCGTGTCTAGCTTAATGAAGTAGGATGTTGGTATTGGTGACACTTGA	848	

Db	917	CGAATGTGGGTCGCCTATTGTCAAAAGTGGCGAAATGTGCGGAAGTTACCAACCAC	976
OY	849	GAAAAGTGACAACATGCGGAGSTTCAAGACTTGATATGCGCACACANMATGCTGAG	908
Db	977	GTGGGGGTGAACCCATGCCCAAGTACACGGTGGCGGATGGCGGTAAAAAACCCAAG	1036
OY	909	GTGGACAGATTATTAATACGTTAGAGCATGTAGCATTTTGGTTCAACACCACCTGTG	968
Db	1037	GTGGACCGCCTATGTCACTACATCGAACCCCTBTAGCCATCATGTGTACGCCACCGTCT	1096
OY	969	TTAACAAATTTATTGATTGTAGAATAGCANNAAGT	1005
Db	1097	GTGACGACATCATCGCCGCTGGCGMAGCGCGCTGTGT	1133
RESULT	15		
AEO00148			
LOCUS			
DEFINITION	AE000148	11692 bp DNA BCT	12-NOV-1998
ACCESSION	AE000148		
VERSION	AE000148.1		
KEYWORDS	GI:1786614		
SOURCE			
ORGANISM	Escherichia coli.		
REFERENCE	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.		
AUTHORS	Blattner,F.R., Plunkett,G. III, Bloch,C.A., Perna,N.T., Burland,V., Riley,M., Coliado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F., Gregor,J., Davis,N.W., Kirkpatrick,H.A., Goeden,M.A., Rose,D.J., Mau,B. and Shao,Y.		
TITLE	The complete genome sequence of Escherichia coli K-12		
JOURNAL	Science 277 (5331), 1453-1474 (1997)		
MEDLINE	97426617		
REFERENCE	2 (bases 1 to 11692)		
AUTHORS	Blattner,F.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.		
COMMENT	Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459		
REFERENCE	3 (bases 1 to 11692)		
AUTHORS	Blattner,F.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.		
COMMENT	Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459		
REFERENCE	4 (bases 1 to 11692)		
AUTHORS	Plunkett,G. III.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-OCT-1998) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA		
COMMENT	This sequence was determined by the E. coli Genome Project at the University of Wisconsin-Madison (Frederick R. Blattner, director). Supported by NIH grants HG00301 and HG01428 (from the Human Genome Project and NCHGR). The entire sequence was independently determined from E. coli K-12 strain MG1655. Predicted open reading frames were determined using Genemark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA, 30332 [e-mail: mark@amber.gatech.edu]. Open reading frames that have been correlated with genetic loci are being annotated with CG Site Nos., unique ID nos. for the genes in the E. coli Genetic Stock Center (CGSC) database at Yale University, kindly supplied by Mary Berlyn. A public version of the database is accessible (http://cgsc.biology.yale.edu). Annotation of the genome is an ongoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site (http://www.genetics.wisc.edu). ***The E. coli K-12 sequence and its annotations are periodically updated; this is version M54. No		

sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Monica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.

```

source
1. 11692
/location/Qualifiers
/organism="Escherichia coli"
/strain="K-12"
/sub_strain="MG1655"
/db_xref="taxon:562"

promoter
17. .46
/note="factor Sigma70; predicted +1 start at 432203"
complement(38. .67)
/note="factor Sigma70; predicted +1 start at 432181"

gene
76. .525
/gene="ybaD"
/note="b0413"
76. .525
/gene="ybaD"
/function="orf, hypothetical protein"
/note="0149; 100 pct identical to YBA_ECOLI SW: P25538"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAC73516.1"
/db_xref="gi:178615"
/db_xref="PID:91786615"
/transl_table="MHCPFCPAVDKTVKIDSRVLVGGSSVRRRQCLVCNERTFTFEVALVLMRVKSNVDREPEENKRLSCMLALKRPVSSDVEMALNHKLSQLRATGEREEVPSKIGVLVMDQLKDKVAVIRASVYRSFEDIKERGEIARLEQ"
complement(88. .116)
/note="factor Sigma70; predicted +1 start at 432231"
418. .446
/gene="ybaD"
/note="factor Sigma32; predicted +1 start at 432603"
529. .1632
/gene="ribD"
/note="b0414"
529. .1632
/gene="ribD"
/function="enzyme; Biosynthesis of cofactors, carriers: Riboflavin"
/note="0367; 100 pct identical to RIBB_ECOLI SW: P25539; alternate gene name ribG"
/codon_start=1
/transl_table=11
/product="bifunctional pyrimidine deaminase/reductase in pathway of riboflavin synthesis"
/protein_id="AAC73517.1"
/db_xref="gi:178616"
/db_xref="PID:91786616"
/transl_table="MODEYVMAKALKAORGRFTTHPNPNCVIVKDGELVGBYHQ RAGEBVAHLRMAGKERAGTAVTLEPCSHGRPPCCDALIAGVAVVMSMOD PNPQVAGGIVRLIOAGIDVSHGLMSFAEOLNKGFLKRMRTGFPYIOIKGALSIDR TMAGSGEOWITSPQARDVOLLRAQSAIITSSSTVYIADPALTVRSELEDEOTALI TPQVLRQPIRVIDSQNRVPRHIVQOPEETWARTQEDSRMPETVRLLIPEKRGHLDIVLMQIDKQOINSIVVEAGPTIAGALLAGVDELIVYIAPKLGSIDRGLCL TLPGLEKADAPQEFKEIRHVPDVCILHVG"
1543. 1571
/gene="ribD"
/note="factor Sigma32; predicted +1 start at 433728"
1721. .2191
/gene="ribH"
/note="b0415"
1721. .2191
/gene="ribH"
/EC_number="2.5.1.9"
/function="enzyme; biosynthesis of cofactors, carriers: Riboflavin"

```

4

```

/note="0156: 100 pct identical to R15B_ECOLI SW:
P25340/residues"
/codon_start=1
/transl_table=11
/product="riboflavin synthase, beta chain"
/protein_id="AAC73518.1"
/db_xref="GI:1786617"
/db_xref="PID:g1786617"
/translation="MNINENVAVTPPARVAITARENNFINDSLEGALDALKRGQV
KIDENTVAVPGAEVLGALGAKGKDAVALATVIRGTAHFEEVAGASGMLA
HVAQSELPVAFGLVTRESIEQAIERRAGTKAGKGAELTALFEMINVLKAIKA"
2116 2133
/feature="r1bH"
/note="Factor Sigma54; predicted +1 start at 434290"
2211 2630
/feature="nusB"
/note="P0416"
2211 2630
/feature="nusB"
/function="factor: RNA synthesis, modification, DNA
transcription"
/note="0133: 100 pct identical to NDSB_ECOLI SW: P04381;
alternate gene name sayB"
/codon_start=1
/transl_table=11
/product="transcription termination: L factor"
/protein_id="AAC73519.1"
/db_xref="GI:1786618"
/db_xref="PID:g1786618"
/translation="MKRAARRARBECAVOALYSWQSDINDVDEQFLAEQDKVDVD
VIRFELLAGATNTATYLDKGLPYSRLLEEQYKRVNLALIELESKSDVPYKVK
AINEALELAKSFGEDSHKTFVNGVLKAPVIRPNKK"
2642 2668
/note="factor sigma70; predicted +1 start at 434825"
2650 2687
/note="factor sigma70; predicted +1 start at 434844"
2679 2706
/note="factor sigma70; predicted +1 start at 434863"
2708 3685
/feature="th1L"
/note="P0417"
2708 3685
/feature="th1L"
/function="enzyme: Biosynthesis of cofactors, carriers:
thiamin"
/note="0325: 56 pct identical to 273 aa of GB:
HU32756_3ACCESSION: U32756 (345 aa)"
/codon_start=1
/transl_table=11
/product="thiamin-monophosphate kinase"
/protein_id="AAC73520.1"
/db_xref="GI:1786619"
/db_xref="PID:g1786619"
/translation="MACEGSEAFEDYRRSSRLDVELGICDDCALINPEKQTLIAH
STDIIVAGNHEFLPDIPADILAKALAVNLSDLAKMADKADMTALTILPDIIDEMLEAH
FSDSLFDLNLNYDMQLIGDGTGSPLSMTKILRGFVPMGRALTRSGAPGWIIVTGG
PDSGAAGAIQLQNLQVADKADKADYILKRLRSPSILLOGALRDLANSAIDSLDGLG
SDLGIIVASQSGRIIDLALPSPDALSRHVEPQALFNALSAGEDYELCTPELNLNR"
GALVADLGHLEVPPTICGQNTADIEGLCFIRQEPYTLDMKGYDHFAFP"
3663 4181
/feature="pgpA"
/note="P0418"
3663 4181
/feature="pgpA"
/function="enzyme: Macromolecule synthesis, modification:
phospholipids"
/note="017: 50 pct identical (2 gaps) to 155 residues of
hypothetical protein YD06_HAEIN SW: P44157 (163 aa)"
/codon_start=1
/transl_table=11
/product="phosphatidylylglycerophosphatase"

```

Best Local Similarity 52.3%; Pred. No. 0.0024;
Matches 145; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

```

OY 729 AATTATGATTATGCGATTCACCTTCCAAATATGCTACAGGTCANACAGGTGTTATC 788
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 542 ATTACATGCGCGCGCGCTAAAGCTGGCGCAACGAGGAGCTTTTACCACGATCCCAACC 601
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 789 CACCGGTGGCGCGCTGTGTAGTTAATGAGTAGATGTTGTGTATTGTCACACTTGA 848
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 602 CGAATGTCGGGTGCGTCATTGTCAAGATGGCAAAATGTGCGTGAAGTTACCAAC 661
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 849 GAAAAGGTGACAGCATGGGAGGTTCAAGCACTGTATGCGACACANAAATGCTGAAG 908
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 662 GTGCGGTGAACCAACATGCGCAAGTACACGCGTGGGTATGCGGGTGAAAAAGCCAAAG 721
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 909 GTGCGACGATTATATTAGCTAGAGCCATGTAGTCATTTGGTCAACACCACTGTG 968
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 722 GTGCGACCGCCTATGTCACTGCAACCTGTAGCCATCATGTGTCAGCCACCGTGT 781
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 969 TTAACAAATTTATGTGTAGATAGCANAAAGTAGT 1005
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 782 GTGACGCACTCATCGCGCGCTGGCGTAGCGCGGTGT 818
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: November 19, 2000, 04:26:04
Job time: 12947 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 05:07:08 ; Search time 118.16 Seconds
(without alignments)
4034.495 Million cell updates/sec

Title: US-08-978-456-3
Perfect score: 1269
Sequence: 1 AANACCAATCCATTCGGA.....GTTAACAGATGCTATAG 1269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID56/gcgdata/geneseq/NA1980.DAT:*
2: /SID56/gcgdata/geneseq/NA1981.DAT:*
3: /SID56/gcgdata/geneseq/NA1982.DAT:*
4: /SID56/gcgdata/geneseq/NA1983.DAT:*
5: /SID56/gcgdata/geneseq/NA1984.DAT:*
6: /SID56/gcgdata/geneseq/NA1985.DAT:*
7: /SID56/gcgdata/geneseq/NA1986.DAT:*
8: /SID56/gcgdata/geneseq/NA1987.DAT:*
9: /SID56/gcgdata/geneseq/NA1988.DAT:*
10: /SID56/gcgdata/geneseq/NA1989.DAT:*
11: /SID56/gcgdata/geneseq/NA1990.DAT:*
12: /SID56/gcgdata/geneseq/NA1991.DAT:*
13: /SID56/gcgdata/geneseq/NA1992.DAT:*
14: /SID56/gcgdata/geneseq/NA1993.DAT:*
15: /SID56/gcgdata/geneseq/NA1994.DAT:*
16: /SID56/gcgdata/geneseq/NA1995.DAT:*
17: /SID56/gcgdata/geneseq/NA1996.DAT:*
18: /SID56/gcgdata/geneseq/NA1997.DAT:*
19: /SID56/gcgdata/geneseq/NA1998.DAT:*
20: /SID56/gcgdata/geneseq/NA1999.DAT:*
21: /SID56/gcgdata/geneseq/NA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1250	98.5	1269	18	DN-encoding-a-Rib
2	1250	98.5	1269	18	Pyrimidine-deamina
3	897.6	70.7	7586	18	Staphylococcus aur
4	352.2	27.8	1029	20	Pyrimidine deamina
5	352.2	27.8	3336	20	Riboflavin biosynt
6	352.2	27.8	3336	20	S. aureus riboflav
7	352.2	27.8	3336	20	Riboflavin biosynt
8	352.2	27.8	3336	20	S. aureus riboflav
9	229	18.0	501	18	DNA encoding a Sta
10	172.8	13.6	223	18	Staphylococcus aur
11	103	8.1	5567	12	B. subtilis rib (ri
12	103	8.1	5567	20	B. subtilis rib op

C	13	103	8.1	5567	20	X81948	B. subtilis rib op
	14	103	8.1	5567	20	V71600	Bacillus subtilis
	15	103	8.1	5567	21	Z93752	Rib operon of Baci
	16	100.4	7.9	7481	18	V74547	Staphylococcus aur
	17	80.4	6.3	1101	20	X77179	S. pneumoniae ribg
	18	80.4	6.3	3417	20	X77183	S. pneumoniae ribg
	19	80.4	6.3	3417	20	X77574	S. pneumoniae ribg
	20	80.4	6.3	3417	20	X79558	S. pneumoniae ribg
	21	80.4	6.3	3417	20	V72454	S. pneumoniae ribg
C	22	80.4	6.3	14273	19	V52173	Streptococcus pneu
	23	64.4	5.1	1230	20	X82287	APP ribg gene sequ
	24	64.4	5.1	4312	19	V27196	Actinobacillus ple
	25	64.4	5.1	4312	20	X82289	Nucleotide sequenc
	26	60.8	4.8	546	19	Z96259	S. pneumoniae deri
C	27	48	3.8	2418	13	Q27886	P. falciparum GBP13
	28	45	3.5	1864	8	N71405	Sequence of ANS-1
	29	43.6	3.4	11309	20	X20250	Borrelia burgdorfe
	30	42.6	3.4	1206	21	A27366	Sinorhizobium meli
C	31	42.6	3.4	4248	21	A27364	Sinorhizobium meli
	32	42.4	3.3	5830	20	X13218	Enterococcus faeca
	33	42	3.3	4968	21	Z55696	Heliothis armigera
	34	41.6	3.3	19124	18	T72882	Plasmodium var-7 g
	35	41.6	3.3	19124	21	Z98287	Plasmodium var-7 p
	36	41.4	3.3	1864	15	Q78952	Aspergillus nidula
	37	41.2	3.2	10502	20	X20258	Borrelia burgdorfe
C	38	40.6	3.2	1463	20	X13655	Enterococcus faeca
	39	40.2	3.2	1444	21	Z94432	Plasmodium falcipa
C	40	39.6	3.1	9834	18	V74348	Staphylococcus aur
	41	39.4	3.1	2033	19	Z26377	DNA encoding Helic
	42	39	3.1	14752	20	X20256	Borrelia burgdorfe
	43	38.6	3.0	1187	18	V74711	Staphylococcus aur
	44	38.2	3.0	783	11	Q04064	Fowl Pox Virus pro
C	45	38.2	3.0	26776	20	X20254	Borrelia burgdorfe

ALIGNMENTS

RESULT	1
1	
TS4023	
ID	TS4023 standard; DNA: 1269 BP.
XX	
AC	TS4023:
XX	
DT	01-SEP-1998 (first entry)
DE	DNA encoding a Riboflavin biosynthesis protein RIBG.
XX	
KW	Staphylococcus aureus protein; ribozyme; antisense sequence; control;
KW	Staphylococcal gene; regulatory element; bacterial gene expression;
KW	vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;
KW	toxic shock syndrome; ss.
XX	
OS	Staphylococcus aureus.
XX	
FH	
FT	Key Location/Qualifiers
FT	CDS 734..1033
FT	/*tag= a
XX	
XX	
XX	WO9730070-A1.
XX	
XX	PD 21-AUG-1997.
XX	
PF	19-FEB-1997: 97WO-US02318.
XX	
PR	20-FEB-1996: 96US-0011888.
XX	
PA	(SMRK) SMITHKLINE BEECHAM CORP.
XX	
XX	Black MT, Burnham MK, Hodgson JE, Knowles DIC, Nicholas RO;
PI	Pratt JM, Reichard RW, Rosenberg M, Ward JM;
XX	
DR	WPI: 1997-424969/39.

DR P-PSDB; W28084.
 XX Novel polypeptide(s) from *Staphylococcus aureus* strain WCUH29 - used
 PT to isolate antimicrobial compounds, and in vaccines against *S.*
 PT aureus infection
 XX
 PS Claim 9; Page 825; 989pp; English.
 XX
 CC The present sequence encodes a *Staphylococcus aureus* protein, that,
 CC based on homology with an *Bacillus subtilis* protein, is believed to be a
 CC Riboflavin biosynthesis protein RibG. The present sequence was obtained
 CC from a library of clones of *S. aureus* WCUH 29 in *Escherichia coli*. The
 CC DNA sequence can be used in the construction of ribozymes and antisense
 CC sequences to control the expression of *Staphylococcal* genes. The DNA
 CC sequence is also useful as a source of regulatory elements for the
 CC control of bacterial gene expression. The encoded protein may be used
 CC to produce vaccines to enable a host to produce specific antibodies
 CC with antibacterial action. These vaccines and antibodies would protect
 CC a host against invasion by *S. aureus*, and conditions relating to
 CC *Staphylococcal* infection, e.g. *Staphylococcal* food poisoning, scaled
 CC skin syndrome, and toxic shock syndrome.
 XX
 SQ Sequence 1269 BP; 435 A; 188 C; 256 G; 371 T; 19 other;
 Query Match: 98.5%; Score 1250; DB 18; Length 1269;
 Best Local Similarity 100.0%; Pred. No. 6.7e-291;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 661 ATTTTCTCCTTGGACATCTTAATTCATGATGAGGATTTTGTATTAGAGGTATCA 720
 DB 661 attttctccttggacatttaattcatgatgaggttttgttattagaggtatca 720
 QY 721 TTTAGTCATTTATGATTTATGCGATTTCACCTTCCAAATATGTTACAGGTANNCAGG 780
 DB 721 tttagtcattttagatattatgctgatttcacattccaaatattgttacaggtanncagg 780
 QY 781 TGTATTCACCCCGTTGGCGCTGTGTAGTTAATGAAGTAGAGTGTGTATTTGGTGC 840
 DB 781 tgtattcaccccgttggcgctgtgtagttaatgaagttagagtggttatgttggctgc 840
 QY 841 ACACCTTGAGAAAAGGTGACAGCATGCGGAGTTCAAGCACTTGATATGGCACACANAA 900
 DB 841 acaacttgagaaaaggtgacagcatgcgaggttcaagcacttgatattggcacacanaa 900
 QY 901 TGTGGAAGGTGCGACATTTATATTTAGCTTAGAGCCATGTAGTCATTTTGGTCAACACC 960
 DB 901 tgtggaaggtgcgacatttatattacgttagagccatgtagtcattttgttcaacacc 960
 QY 961 ACCCTGTATACAAATTTATTTGATTGATAGATAGCANAGTAGTATTACNCAANAAG 1020
 DB 961 accctgtatacaaatttatgttattgattagatagcanagtagtatttacncaanaag 1020
 QY 1021 ACAATTCGTTAGACACATGGTGATGAGAGCTTACGGGGCTCCAGGTATTGTAGAGG 1080
 DB 1021 acaatttcgttagacacatggtgatgagagcttaccggggctccaggtattgttagagg 1080
 QY 1081 TTGATTTGCGTTGGATGATATACAGGGCATCACAATTTATACCAAGACTTTTAAAGCA 1140
 DB 1081 ttgatttgctgttgatgatatacagggcatcacaaatttataccaagactttttaaagca 1140
 QY 1141 AAAAGCAAGCAACTTGCCCAAAATATACAGTGAAGTNTCTTGAAGTTAGATGAGTA 1200
 DB 1141 aaaagcaagcaacttgcccaaaatatacagtgaaagtntcttgaagtttagatgagta 1200
 QY 1201 AACAAAGCGAATTGATATGAGCAAAAGTCAATGATTACTTACAAAGAGGTATTAACAGA 1260
 DB 1201 aacaaagcgaaattgatattagcaaaaagtcaattgattacttaacaaagaggttaacaga 1260
 QY 1261 TGTCTATAG 1269
 DB 1261 tgtctatag 1269

RESULT 2
 X59910
 ID X59910 standard; DNA; 1269 BP.
 XX
 AC X59910;
 DT 04-AUG-1999 (first entry)
 XX
 DE Pyrimidine deaminase and pyrimidine reductase (ribG) ORF sequence.
 XX
 KW Pyrimidine deaminase; pyrimidine reductase; ribG; antagonist;
 KW agonist; antimicrobial; antibiotic; Helicobacter pylori infection;
 KW H. pylori-induced cancer; gastrointestinal carcinoma; gastric ulcer;
 KW gastritis; ss.
 XX
 OS *Staphylococcus aureus*.
 XX
 PN W09926475-A1.
 XX
 PD 03-JUN-1999.
 XX
 PF 23-NOV-1998; 98WO-0524857.
 XX
 PR 25-NOV-1997; 97US-0978456.
 XX
 PA (SMIR) SMITHKLINE BEECHAM CORP.
 XX
 PI Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DJC;

PI Lonetto MA, Nicholas RO, Palmer LM, Pratt JM, Reichard RW;
PI Rosenberg M, Trani CM, Ward JM, Warren RL;
xx
DR WPI. 1999-347572/29.
DR P-PSDB: Y15914.
xx
PT New *Staphylococcus aureus* polypeptide and polynucleotide useful in
xx the treatment of gastric ulcer and gastritis
xx
PS Claim 24; Page 5-6; 48pp: English.
xx
CC The present sequence encodes a *Staphylococcus aureus* pyrimidine
CC deaminase and pyrimidine reductase (ribG) polypeptide. *Staphylococcus*
CC *aureus* ribG and its antagonists are used to treat individuals in
CC need of them. Disease related to expression or activity of ribG can
CC be determined by analysing the nucleic acid sequence encoding ribG
CC or detecting the ribG polypeptide in a sample. ribG can also be used
CC to identify antagonists or agonists. ribG, or its related nucleic acid,
CC also has use as a vaccine to induce an immunological response in an
CC animal. Antimicrobial compounds (e.g. agonists and antagonists of ribG),
CC especially broad-spectrum antibiotics, may be of use in the treatment
CC of *Helicobacter pylori* infection. This should decrease the advent of
CC *H. pylori*-induced cancers, such as gastrointestinal carcinoma. The
CC treatment should also cure gastric ulcers and gastritis.
xx
SQ Sequence 1269 BP: 435 A; 188 C; 256 G; 371 T; 19 other;

Query Match 98.5%; Score 1250; DB 20; Length 1269;
Best Local Similarity 100.0%; Pred. No. 6,7e-291;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AANCAACCAATCCNATTGGAGAGNAATCCAAATCAATCCCGANCCCAATCCAAATTA 60
DB 1 aancaccaatccnatggagagnaaaccacaatcnaaccgganmcccaatccaaatlaa 60
OY 61 TTAAGTCCAAAGTTTGGACATTACCAATATGATTCGATGAGGTCAATGNCANCG 120
DB 61 ttaagtccaaagtttggacaattaccacaatgatgccgatgaggtccaatgncancng 120
OY 121 GTGTTAATAACTAGCAAAATGNTGNAATAGATAGTAGANCAAGTTGGCATACAGTNT 180
DB 121 gttgtaataactagcaaaatgntgnaatagatagtagancaagttggcatacagtnt 180
OY 181 CTCNATTAATATGCTTATGAAATGCAATGACGACAAACAGCGCAATTCATCAATAT 240
DB 181 ctcnatataatgcttattgaaatgcaatgacgacaaacagcgcaattcatcaatatt 240
OY 241 TTGTACTAGAGATATTAAGAAAGAGCTCTATAGAGCAATTAAGTTGATTTTA 300
DB 241 ttgtaactagaagataataaagaaagcctctatagaagcgaattgaagttgatatta 300
OY 301 ATGCTGTAGTAGAATCATATCAATGAGATGCCATATAGTACTCAGATTATTAATTA 360
DB 301 atgctgttagtagaataatcatatcaatgagatgacctatagttactcagattataatt 360
OY 361 AAAACCGTATTAATGTTTTTTAGAAAACATATGTCATTTTAATGTAAGTGA 420
DB 361 aaaacggtatataatgttttttagaaaacatatagtatcatltaaatgttagtga 420
OY 421 TACTAGTCTCAAAATATCTAATACATTTATATATTTCTTCCGGGCGAGGTGAA 480
DB 421 tactagtgctcaaaatatactatacaattcatatataatcttcg99gcag99tgaa 480
OY 481 ATTCCCAACCGGACAGTAATAAAGCCCTGACCTGCTAATATGTTTCAATATTAGTG 540
DB 481 attcccaacggcagtaataaagcctgacctgactaataatgtttcatattagtgctg 540
OY 541 AACTAGTGAATTTAGACCCGACACTTAAAGTCTGATGGGGAAGAAAGATTTAAT 600
DB 541 aactagtgaatttagaccgacactttaaagtctgatggggaagaaagttaattat 600
OY 601 GCACAAGATATGTAAGCTATTGTAATAAATGTGTACAAATAGCTTATTTAAGCAT 660

DB 601 gcacaagataatgtagcgtattgttaaaatgtgtacaatagcttataacgata 660
OY 661 ATTTTTCCTTCCTTGATCTTAATTCATGATGAGAGATTTTGTGTTATAGAGTGATCA 720
DB 661 atttttcccttccttgatcttaatttcattgatgagagattttgtttataaggtgata 720
OY 721 TTGAGTCAATTAATGATTTATGATTCATCACTTCCAAATATGTAAGGCAACAGG 780
DB 721 ttgagtcatttaataatgattatgcatccaacttccaatatgtacaaagtccaacagg 780
OY 781 TGTATATCCACCCGTTGGCGCTGTTGTAATGTAAGGTAGAGATGTTGTTGTTGC 840
DB 781 tgtatattccaccggttgccgctgttgtaattgtaaggtagagattgttgattgtgc 840
OY 841 ACACTTGAGAAAAGTGCACAGCATGCGGAGGTTAAGCAGCTGATATGGCCACANAA 900
DB 841 acacttgagaaaagtgacacagcatgcgaggttcaagcacttgatagycacaacanaa 900
OY 901 TGCTGAGGTGCGACATTATATACGTTAGAGCCATGTACTCATTTTGGTTCAACAC 960
DB 901 tgctgaggtgcgacatttatatcagttagagccatgtactcatTTTGGTTCAACAC 960
OY 961 ACCCTGTGTTAACAAAATTTATGATTGTAAGATAGCANAAAGTAGATTACNCAACANA 1020
DB 961 accctgtgtaacaaaatttatgattgtaagatagcanaaagtagattacncaacanaa 1020
OY 1021 ACAATTCGTTGACACACATGGGTGATGACGCTTACGGGGCTCCAGGTTATTGAGG 1080
DB 1021 acaatttcgttgacacacatgggtgatgacgcttaccgggctccaggttattgagg 1080
OY 1081 TTGAATTTGGTTGATGATGAAGCGGCATCACATTAATCCAAAGACTTTTAAAGA 1140
DB 1081 ttgaatttggttgatgatgaagcggcatcacatttaataccaaagactttttaaaga 1140
OY 1141 AAAAGCAAAACACTTGGCCACAAATTAACAGTAAGTAAGTCTTGAAGTTTATGATGG 1200
DB 1141 aaaagcaaaacacttggccacaaatttaacagtaagtaagtcttgaagtttattgatgg 1200
OY 1201 AACAAAGCAATTTGATTAATGAGCAAGTCATGATTTACTAACAAGAGTTAAACAGA 1260
DB 1201 aacaaagcaatttgattaatgagcaagtcattgattacttaacaagaagttaacaea 1260
OY 1261 TGCTATAG 1269
DB 1261 tgcataag 1269

RESULT 3
V74453/C
ID V74453 standard; DNA: 7588 BP.
XX
AC V74453:
XX
DT 16-MAR-1999 (first entry)
XX
DE *Staphylococcus aureus* contlg SEQ ID #142.
XX
KW Computer readable medium; vaccine; *S. aureus* infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.
XX
OS *Staphylococcus aureus*.
XX
FH Key Location/Qualifiers
FT misc-feature 661..720
FT misc-feature 721..780
FT misc-feature 781..840
FT misc-feature 841..900
FT misc-feature 901..960
FT misc-feature 961..1020
FT misc-feature 1021..1080
FT misc-feature 1081..1140
FT misc-feature 1141..1200
FT misc-feature 1201..1260
FT misc-feature 1261..2520

```
FT      /tag- b
FT      /note- "these bases represent a line of missing text in
FT      the sequence listing in the specification. They
FT      are included to maintain the nucleotide numbering
FT      given in the specification for this DNA sequence"
FT      misc-feature
FT      4261..4320
FT      /tag- c
FT      /note- "these bases represent a line of missing text in
FT      the sequence listing in the specification. They
FT      are included to maintain the nucleotide numbering
FT      given in the specification for this DNA sequence"
FT      misc-feature
FT      6061..6120
FT      /tag- d
FT      /note- "these bases represent a line of missing text in
FT      the sequence listing in the specification. They
FT      are included to maintain the nucleotide numbering
FT      given in the specification for this DNA sequence"
FT      EP786519-A2.
FT      30-JUL-1997.
FT      07-JAN-1997: 97EP-0100117.
FT      05-JAN-1996: 96US-0009861.
FT      (HUMA-) HUMAN GENOME SCI INC.
FT      Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,
FT      Rosen CA;
FT      WPI; 1997-374922/35.
FT      Polynucleotide(s) and proteins derived from Staphylococcus aureus
FT      stored on computer readable medium and used in the production of
FT      anti-S.aureus vaccines
FT      Claim 1; Page 740-744; 3271pp; English.
FT      This sequence represents one of 5191 Staphylococcus aureus DNA sequences
FT      of the invention. The DNA sequences are recorded on a computer readable
FT      medium, preferably selected from a floppy or hard disk, random access
FT      memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
FT      the S.aureus DNA sequences allows putative functions to be assigned so
FT      that protein-encoding or regulatory regions of commercial, therapeutic or
FT      industrial importance can be obtained. Specifically, sequences which are
FT      likely to encode antigens have been identified and these polypeptides can
FT      be used in a vaccine composition against S.aureus infection. The
FT      polypeptides can also be used in a kit for the immunodetection of
FT      S.aureus in a sample. S.aureus is implicated in numerous human diseases,
FT      including cellulitis, eyelid infections, food poisoning, osteomyelitis,
FT      skin and surgical wound infections, scalded skin syndrome, toxic shock
FT      syndrome, etc. Organisms transformed with the DNA sequences can be used
FT      for recombinant production of the polypeptides. The new DNA sequences
FT      (and their fragments) are useful as primers or probes for isolating
FT      homologues of any of the S.aureus DNA sequences contained on the
FT      computer readable medium.
FT      Sequence 7588 BP; 2333 A; 1341 C; 1061 G; 2613 T; 240 other;
SQ
Query Match 70.7%; Score 897.6; DB 18; Length 7588;
Best Local Similarity 88.1%; Pred. No. 3.7e-206;
Matches 1058; Conservative 0; Mismatches 122; Indels 21; Gaps 11;
```

```
QY      188 ATATGATGCTTTAGATCGAATGAGCAACAGCCAGTACATCAATTAATTTGTAC 247
D      6393 ATATGATGCTTTAGATCGAATGAGCAACAGCCAGTACATCAATTAATTTGTAC 6334
QY      248 TAGAGATTAATAAGAAAGCGCTCTATAGAGCAAGTTGAAGTTGATTTAATGTCG 307
D      6333 TAGAGATTAATAAGAAAGCGCTCTATAGAGCAAGTTGAAGTTGATTTAATGTCG 6274
QY      308 TTGATAGAAATCATATCAATGAGATGCTATAGTACTGATTAATTAATTAACCG 367
D      6273 TTGATAGAAATCATATCAATGAGATGCTATAGTACTGATTAATTAATTAACCG 6214
QY      368 TCATTAATTTGTTTTAGAAAACATATAGTATCATTTTAAATGATGACATTAAGTACG 427
D      6213 TCATTAATTTGTTTTAGAAAACATATAGTATCATTTTAAATGATGACATTAAGTACG 6154
QY      428 TACTCAAAATATCATATACAAATTTCAATATAATTTCTTTGGGGCAGGGTGAATTCGA 487
D      6153 TACTCAAAATATCATATACAAATTTCAATATAATTTCTTTGGGGCAGGGTGAATTCGA 6094
QY      488 ACCGGCAGTAAATAAAGCCGCGACCTGCTAATATGTTTCATTTAGTGGCTGATCTAGT 547
D      6093 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6034
QY      548 GAGATTTCTAGAGCCGACAGTAAAGTCTGGATGGAGAAAAGAAATGTAATTCGACAAA 607
D      6033 GAGATTTCTAGAGCCGACAGTAAAGTCTGGATGGAGAAAAGAAATGTAATTCGACAAA 5974
QY      608 GATTAATGTAGCGTATTTGTAATAAATGTGACAAATAGGCTTATTTAAGCATTAATTTTC 667
D      5973 GATTAATGTAGCGTATTTGTAATAAATGTGACAAATAGGCTTATTTAAGCATTAATTTTC 5914
QY      668 TCCTTTGATCTTAATTCATGATGATGAGATTTTGTGTTATAGAGTATCATTTGAGT 727
D      5913 TCC-TTGCATCTTAATTCATGATGATGAGATTTTGTGTTATAGAGTATCATTTGAGT 5856
QY      728 CAATTTATGATATGATGATTCACCTTCCAAATATGATGTAAGGTCANACAGGTGTTAAT 787
D      5855 CAATTTATGATATGATGATTCACCTTCCAAATATGATGTAAGGTCANACAGGTGTTAAT 5796
QY      788 CCACCCGTTGGCGCTGTTGATTAATGAAGTAGAGATTTGTTGTTAGTGGTCACACTGG 847
D      5795 CCACCCGTTGGCGCTGTTGATTAATGAAGTAGAGATTTGTTGTTAGTGGTCACACTGG 5736
QY      848 AGAAAGGTGACAGATGCGGAGGTTCAAGCCTTGATATGGCAACAAMATGCTGAA 907
D      5735 AGAAAGGTGACAGATGCGGAGGTTCAAGCCTTGATATGGCAACAAMATGCTGAA 5676
QY      908 GGTGCGACGATTTATATATGATGATAGAGCATGATGATTTTGGTCAACACACCCTGT 967
D      5675 GGTGCGACGATTTATATATGATGATAGAGCATGATGATTTTGGTCAACACACCCTGT 5616
QY      968 GTTAACAAATTTATGATTTGAAGATAGACANAAGTATTAACNCAACANAAGCAATTC 1027
D      5615 GTTAACAAATTTATGATTTGAAGATAGACANAAGTATTAACNCAACANAAGCAATTC 5556
QY      1028 CGTTAGACACACATGGGTATGAGAGCTTACGGGGCTCCACGGTATTTAGGGTGAAT 1087
D      5557 CGTTAGACACACAT-GGTATGAGAGCTTACGGGGCTCCACGGTATTTAGGGTGAAT 5499
QY      1088 GCGTGTGATGATGAGCGGATCACAAATTAACCAAGACTTTTAAAGCAAAAAGCA 1147
D      5498 -----GATGATTAAGCGGATCACAAATTAACCA--AGACTTTTAAAGCAAAAAGCA 5448
QY      1148 AAGCAACTTGCACAAAATTTACAGTGAAGTNTCTTGAAGTTTACAGGTGAACAAAG 1207
D      5447 AAGCAACT--GCCACAAATTTACAGTGAAGTATC--TGCAAGTTTACAGGTGAACAAAG 5391
QY      1208 CGAATGATATAGTACAAAGTCAATGATTAATTAACAAAGAGTTTAAACAAGATGCTAT 1267
D      5390 ---AATGATTAATAGTACAAAGTCAATGATTAATTAACAAAGAGTTTAAACAAGATGCTAT 5334
QY      1268 A 1268
```

Db 5333 A 5333

RESULT 4

ID X59909 standard; DNA; 1029 BP.

XX X59909;

DT 04-AUG-1999 (first entry)

XX Pyrimidine deaminase and pyrimidine reductase (ribg) DNA.

XX Pyrimidine deaminase; pyrimidine reductase; ribg; antagonist;
 KM agonist; antimicrobial; antibiotic; Helicobacter pylori infection;
 KM H. pylori-induced cancer; gastrointestinal carcinoma; gastric ulcer;
 KM gastritis; ss.

XX Staphylococcus aureus.

XX Key Location/Qualifiers
 FH 1..1029
 FT /*tag- a
 FT /product= "ribg"

XX MO926475-A1.

XX 03-JUN-1999.

XX 23-NOV-1998; 98MO-US24857.

XX 25-NOV-1997; 97US-0978456.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DJC;

XX Lonetto MA, Nicholas RO, Palmer LM, Pratt JM, Reichard RW;

XX Rosenberg M, Trani CM, Ward JM, Warren RL;

XX MPI: 1999-347572/29.

XX P-PSDB: Y15913.

XX New Staphylococcus aureus polypeptide and polynucleotide useful in

XX the treatment of gastric ulcer and gastritis

XX Claim 6; Page 5; 48pp; English.

XX The present sequence encodes a Staphylococcus aureus pyrimidine
 CC deaminase and pyrimidine reductase (ribg) polypeptide. Staphylococcus
 CC aureus ribg and its antagonists are used to treat individuals in
 CC need of them. Disease related to expression or activity of ribg can
 CC be determined by analysing the nucleic acid sequence encoding ribg
 CC or detecting the ribg polypeptide in a sample. ribg can also be used
 CC to identify antagonists or agonists. ribg, or its related nucleic acid,
 CC also has use as a vaccine to induce an immunological response in an
 CC animal. Antimicrobial compounds (e.g. agonists and antagonists of ribg),
 CC especially broad-spectrum antibiotics, may be of use in the treatment
 CC of Helicobacter pylori infection. This should decrease the advent of
 CC H. pylori-induced cancers, such as gastrointestinal carcinoma. The
 CC treatment should also cure gastric ulcers and gastritis.

XX Sequence 1029 BP; 376 A; 162 C; 198 G; 293 T; 0 other;

XX Query Match 27.8%; Score 352.2; DB 20; Length 1029;

XX Best Local Similarity 88.4%; Pred. NO. 1.4e-75; Gaps 8;
 XX Matches 473; Conservative 0; Mismatches 44; Indels 18;

QY 734 ATGATTATCGATTCACTTCCAAATATGTTACAGGTGCTTAATCCACC 793
 |||||||

Db 1 atgattatcgattcaacttgcacaatagtgtacaaggtcaaacagtgtaataccacc 60

QY 794 GTTGCCGCTGTGTAGTAAAGAGTAGGATTGTGATGTCACATTGAGAAA 853
 |||||||

Db 61 gtggcgctgttgtaagtaagagtagattgtgtatgtgtgacacttgagaaaa 120
 |||||||

QY 854 GGTGACAGCATGGCGAGGTTCAAGCATTGATGCGACACAAANAATGCTGAAGTGGC 913
 |||||||

Db 121 ggtgacaagcatgcggaggttcaagcacttgatagtcacaacaaatgtctgaagtcg 180
 |||||||

QY 914 ACGATTATATTCAGTACGAGCCATGATGATTTGGTTCACACACCCTGGTTAAC 973
 |||||||

Db 181 acgattatatactagtagagccatgtagtcatcttggttcaacacacccgtgttaac 240
 |||||||

QY 974 AAATTATGATTGTAGATAGCAANAAGTAGATTATACCAANAAGACAATTCGGTTAG 1033
 |||||||

Db 241 aaaatttatttgtaagtaagcaaaaagtgtat-taagcaacaaagacatt-cgttag 298
 |||||||

QY 1034 ACACACATGGGTGATGAGACGTTACGGGCTCCACGGTATTGAGGTTAAATTGGCTTG 1093
 |||||||

Db 299 acaacacat-gtgtatgagcgttacggtcacggtatgtagtgatgcgtl----- 351
 |||||||

QY 1094 GATGATGAAGCGGATACAAATATATACCAAGACTTTTAAAGCAAAAGCAAGCAA 1153
 |||||||

Db 352 gatgatgaacgggcatcacattatatacgaacttltt--aaagcaaaagaaagcaa 408
 |||||||

QY 1154 CTTCGCACAAAATTACAGTGAAGTNTCTGAAAATTAGATGGGTAAACGAAGCAATT 1213
 |||||||

Db 409 ct--gccacaattacagtgtaaaagatc-tgcaagtttagatgttaacaaagc---aat 462
 |||||||

QY 1214 GATATGGAACAAAGTCATGATTTACTAACAAAGGTTAAACAGATGCTATA 1268
 |||||||

Db 463 gataatgacaagaatcatgattactacaagaaggttaacaagaatgtcata 517
 |||||||

RESULT 5
 X59913
 ID X59913 standard; DNA; 3336 BP.

XX X59913;

XX 04-AUG-1999 (first entry)

XX Riboflavin biosynthesis operon containing the ribg gene.

XX Pyrimidine deaminase; pyrimidine reductase; ribg; antagonist;

KM agonist; antimicrobial; antibiotic; Helicobacter pylori infection;
 KM H. pylori-induced cancer; gastrointestinal carcinoma; gastric ulcer;
 KM gastritis; ss.

XX Staphylococcus aureus.

XX MO926475-A1.

XX 03-JUN-1999.

XX 23-NOV-1998; 98MO-US24857.

XX 25-NOV-1997; 97US-0978456.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DJC;

XX Lonetto MA, Nicholas RO, Palmer LM, Pratt JM, Reichard RW;

XX Rosenberg M, Trani CM, Ward JM, Warren RL;

XX MPI: 1999-347572/29.

XX New Staphylococcus aureus polypeptide and polynucleotide useful in
 PT the treatment of gastric ulcer and gastritis

PS Example 2; Page 30-31; 48pp; English.

XX The present sequence represents the Staphylococcus aureus
 CC riboflavin biosynthesis operon containing the pyrimidine

CC deaminase and pyrimidine reductase (ribG) gene. *Staphylococcus*
 CC aureus ribG and its antagonists are used to treat individuals in
 CC need of them. Disease related to expression or activity of ribG can
 CC be determined by analysing the nucleic acid sequence encoding ribG.
 CC or detecting the ribG polypeptide in a sample. ribG can also be used
 CC to identify antagonists or agonists. ribG, or its related nucleic acid,
 CC also has use as a vaccine to induce an immunological response in an
 CC animal. Antimicrobial compounds (e.g. agonists and antagonists of ribG),
 CC especially broad-spectrum antibiotics, may be of use in the treatment
 CC of *Helicobacter pylori* infection. This should decrease the advent of
 CC H. pylori-induced cancers, such as gastrointestinal carcinoma. The
 CC treatment should also cure gastric ulcers and gastritis.

CC Sequence 3336 BP; 1202 A; 460 C; 677 G; 997 T; 0 other;

Query Match: 27.8%; Score 352.2; DB 20; Length 3336;
 Best Local Similarity 88.4%; Pred. No. 1.9e-75;
 Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;

```

OY 734 ATGGATTATGCGATTCCAAATATGATGACAAAGTCANACAGGTTTATCCAGCC 793
DB 1 atggattatgcatcattcaacttgcgaataatgcgaagtcacaacagtgcttaccacc 60
OY 794 GTTGGCGCTGTGTAGTTAATGAGGTAGATGTTGTTATGTTGTCACACTTGAGAAA 853
DB 61 gtggcgctgtgttagttaatgaagtagatgttggattgtgtgcacacttgagaaaa 120
OY 854 GGTGACAAAGCATGCGAGGTTCAAGCACTTGATATGCGACANACANATCTGAAGTGG 913
DB 121 ggtgacaagcatgcygaggttcaagcacttgatgacacaacaaatgctgaagttgcg 180
OY 914 ACGATTATATACGTTAGAGCCATGATGATTTGTTGTTCAACACACCGCTGTATAC 973
DB 181 acgattatatacgtttagagccatgtagtcatcttggttcaaccaccctgtgttaac 240
OY 974 AAAATTATGATTGTAGATAGCANAAGTAGTATTACNACANAGACAAATTCGTTAG 1033
DB 241 aaattattgattgttaagtagaagaagaagtagta-tacgacaacaaagacaatt-cgttag 298
OY 1033 ACACATATGGGTGATGAGACGTTACGGGGCTCCACGCTATTGAGGTTGATTCGTTG 1093
DB 299 acaccat-ggtgatgagacgtttagcggctcaggtatgaggttgaagcgtt----- 351
OY 1094 GATGATGAACGGGATCAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCAAA 1153
DB 352 gatgatgaacgggcatcacaaattataccaagaactttt---aaagcaaaagcaagcaa 408
OY 1154 CTGGCCACAAAATTACAGTGAAGTNTCTTGAAGTTTATGATGGGTTAAACAAAGCAATT 1213
DB 409 ct-gccacaatttacagtgaaagatc-tgcaagtttagatgtaaacaaagcg--aat 462
OY 1214 GATAATGCAAAAGTCAATGATTTACTATCAAAAGAGGTTAAACAGATCTCTATA 1268
DB 463 gataatgcaaaagtcataatgattactaacaagaaggttaacaagaatgctcata 517

RESULT 6
X77007
ID X77007 standard; DNA; 3336 BP.
XX AC X77007;
XX XX
DT 06-AUG-1999 (first entry)
XX XX
DE S. aureus riboflavin biosynthesis operon.
XX XX
KM GTP cyclohydrolase II; ribA; diagnosis; vaccine; bacterial adhesion;
KM Helicobacter pylori infection; gastrointestinal carcinoma; gastric ulcer;
KM gastritis; wound healing; infection; drug screening; therapy;
KM riboflavin biosynthesis operon; ss.
XX XX
OS Staphylococcus aureus.

```

XX W09926646-A1.
 PN 03-JUN-1999.
 PD
 XX 24-NOV-1998; 98WO-US25068.
 PF 25-NOV-1997; 97US-0977554.
 PR (SMIK) SMITHKLINE BEECHAM CORP.
 PA Black MT, Burnham MKR, Debouck C, Fedon JC, Hodgson JE;
 PI Jaworski DD, Knowles DJC, Lometto MA, Mooney J, Nicholas RO;
 PI Palmer LM, Pratt JM, Reichard RM, Rosenberg M, Traini CM;
 PI Wang M, Ward JM, Warren RL, Zhong Y;
 DR WPI. 1999-357736/30.
 XX
 XX New *Staphylococcus aureus* ribA polypeptide - and methods for using
 PT the polypeptide to screen for antibacterials
 PT
 PS Claim 24; Page 36-37; 53pp: English.

CC This sequence represents the *S. aureus* riboflavin biosynthesis operon,
 CC which contains the ribA coding sequence of the invention. ribA belongs to
 CC the GTP cyclohydrolase II family. The polynucleotides and polypeptides
 CC may be employed as research reagents and material for the discovery of
 CC treatments and diagnostics for diseases, particularly human diseases.
 CC They can be used for diagnosis of the disease and staging of disease, and
 CC as reagents in differential screening methods. The polynucleotides may be
 CC used as a source for hybridisation probes, and for screening of genetic
 CC mutations, serotype, and identification, and for organism chromosome
 CC identification. The polypeptides can be used to produce antibodies. The
 CC polypeptides can also be used in vaccine formulations, and to identify
 CC agonists and antagonists. These are used to prevent, inhibit or treat
 CC diseases, particularly of *Helicobacter pylori* infections, such as
 CC gastrointestinal carcinoma. Gastric ulcers and gastritis may also be
 CC treated. The polypeptides can also be used to treat wounds and
 CC in-dwelling devices to prevent bacterial adhesion and infection, and to
 CC block ribA protein-mediated mammalian cell invasion. The frequency of
 CC *Staphylococcus* infections has risen dramatically, and it is no longer
 CC common to find isolated *S. aureus* strains that are resistant to standard
 CC antibiotics. The ribA products of the invention can be used screen for
 CC new antibacterial compounds that may target these resistant bacteria.

XX Sequence 3336 BP; 1202 A; 459 C; 678 G; 997 T; 0 other;

Query Match: 27.8%; Score 352.2; DB 20; Length 3336;
 Best Local Similarity 88.4%; Pred. No. 1.9e-75;
 Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;

```

OY 734 ATGGATTATGCGATTCCAAATATGATGACAAAGTCANACAGGTTTATCCAGCC 793
DB 1 atggattatgcatcattcaacttgcgaataatgcgaagtcacaacagtgcttaccacc 60
OY 794 GTTGGCGCTGTGTAGTTAATGAGGTAGATGTTGTTATGTTGTTCAACACTTGAGAAA 853
DB 61 gtggcgctgtgttagttaatgaagtagatgttggattgtgtgtgcacacttgagaaaa 120
OY 854 GGTGACAAAGCATGCGAGGTTCAAGCACTTGATATGCGACANACANATCTGAAGTGG 913
DB 121 ggtgacaagcatgcygaggttcaagcacttgatgacacaacaaatgctgaagttgcg 180
OY 914 ACGATTATATACGTTAGAGCCATGATGATTTGTTGTTCAACACACCGCTGTATAC 973
DB 181 acgattatatacgtttagagccatgtagtcatcttggttcaaccaccctgtgttaac 240
OY 974 AAAATTATGATTGTAGATAGCANAAGTAGTATTACNACANAGACAAATTCGTTAG 1033
DB 241 aaattattgattgttaagtagaagaagaagtagta-tacgacaacaaagacaatt-cgttag 298
OY 1034 ACACATATGGGTGATGAGACGTTACGGGGCTCCACGCTATTGAGGTTGATTCGTTG 1093

```

```
||||| ||||||| |||||||
Db 299 acacacat-gtgtatgagacgttacgggtccacgattatgagtgatgcgtt----- 351
QY 1094 GATGATGAACGGGCATCACATATATACCAAGACTTTTAAAGCAAAAGCAAGCAA 1153
Db 352 gatgatgaacgggcacacacattaccacgaactttt---aaagcaaaagcaagcaa 408
QY 1154 CTGCGCAAAAATTACGTGAAGTNTCTGAAGTTTACATGGGTAAACAAAGCAATT 1213
Db 409 ct--gccacaattacagtgaaagatc-tgcaagtttagatgtaacaaagc---aat 462
QY 1214 GATATATGACAAAGTCATGATTACTTACCAAGAGGTAAACAGATGTCTATA 1268
Db 463 gataatggaacaagtcaatgattactaacaagaaggttaacaagaatgtctata 517

RESULT 7
X60277
ID X60277 standard; DNA; 3336 BP.
XX
AC X60277;
XX
DE 11-AUG-1999 (first entry)
XX
XX Riboflavin biosynthesis operon derived open reading frame.
XX
KM Riboflavin synthase; alpha-subunit; ribB; vaccine;
KM Helicobacter pylori infection; gastrointestinal carcinoma;
KM gastric ulcer; gastritis; bacterial adhesion; infection;
KM cell invasion; bacteriostatic; bactericidal; ss.
XX
OS Staphylococcus aureus.
XX
PN WO9926650-A1.
XX
PD 03-JUN-1999.
XX
PF 24-NOV-1998; 98WO-US25070.
XX
PR 25-NOV-1997; 97US-0977866.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
PI Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DJC;
PI Lonetto MA, Nicholas KO, Palmer LM, Pratt JM, Reichard RW;
PI Rosenberg M, Trani CM, Ward JM, Warren RL;
XX
DR WPI: 1999-370857/31.
XX
PT New Staphylococcus aureus ribB polypeptide, useful in the prevention
PT and treatment of Helicobacter pylori infections
XX
PS Example 2; Page 30-31; 44pp; English.
XX
CC The present sequence represents an open reading frame of the
CC riboflavin biosynthesis operon, containing the ribB sequence. RibB is a
CC riboflavin synthase (alpha-subunit) protein. It is isolated from
CC Staphylococcus aureus. The polynucleotides and polypeptides may be
CC employed as research reagents and material for the discovery of
CC treatments and diagnostics for diseases, particularly human diseases.
CC They can be used for diagnosis of the disease and staging of disease,
CC and as reagents in differential screening methods. The polynucleotides
CC may be used as a source for hybridization probes, and for screening
CC of genetic mutations, serotype, and identification, and for organism
CC chromosome identification. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. These are used to prevent,
CC inhibit or treat diseases, particularly of Helicobacter pylori
CC infections, such as gastrointestinal carcinoma. Gastric ulcers and
CC gastritis may also be treated. The polypeptides can also be used to
CC treat wounds and in-dwelling devices to prevent bacterial adhesion and
CC infection, and to block rib protein-mediated mammalian cell invasion.
CC The ribB polypeptide can also be used to screen for antagonists and
```

```
CC agonists which may be bacteriostatic and/or bactericidal, and so may
CC also be employed to inhibit and treat diseases.
XX
SQ Sequence 3336 BP; 1202 A; 459 C; 678 G; 997 T; 0 other:
Query Match 27.8%; Score 352.2; DB 20; Length 3336;
Best Local Similarity 88.4%; Pred. No. 1.9e-75;
Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;
QY 734 ATGATTTATTCGATTTCACTTCCAAATATGTATCAAGGTCANACAGGTGTAATCCACC 793
Db 1 atgattatcgattcaacttgcaactgtaataatgtaacaagttcaacagtgtaaccacc 60
QY 794 GTTGGCGGTGTAGTTATGAGTAGATGTTGTTAGTGTGTCACACTGAGAAA 853
Db 61 gtgcgcgcgtgtagttagttagttagttagttagttagttagttagttagttagttagt 120
QY 854 GGTGACAGCATGCGAGGTTTCAGACCTTGATATGACACACANANATGCTGAGGTGCG 913
Db 121 ggtgacaagcatgcygaagttcaagcacttgatgtaacaaacaatgctgaagtgcg 180
QY 914 ACGATTTATATTAAGTTAGAGCCATGATGATTTGTTCAACACCACCTGTGTTAAC 973
Db 181 acgattatattacgttagagccatgtagcatttggttcaaccaccctgtgtaac 240
QY 974 AAATATATGATTTGATAGATAGACANAGTAGATTTACNCAACANAGACAAATCCGTTAG 1033
Db 241 aaatattatgattgtagaagaagaagaagtagta-tacgcaacaagaacaatt-cgttag 298
QY 1034 ACACACATGGGTGATGAGACGTTCACGGGCTCCACGATTTGAGGGTTGAATTCGTTG 1093
Db 299 acacacat-gtgtatgagacgttacgggtccacgattatgagtgatgcgtt----- 351
QY 1094 GATGATGAACGGGCATCACATATATACCAAGACTTTTAAAGCAAAAGCAAGCAA 1153
Db 352 gatgatgaacgggcacacacattaccacgaactttt---aaagcaaaagcaagcaa 408
QY 1154 CTGCGCAAAAATTACGTGAAGTNTCTGAAGTTTACATGGGTAAACAAAGCAATT 1213
Db 409 ct--gccacaattacagtgaaagatc-tgcaagtttagatgtaacaaagc---aat 462
QY 1214 GATATATGACAAAGTCATGATTACTTACCAAGAGGTAAACAGATGTCTATA 1268
Db 463 gataatggaacaagtcaatgattactaacaagaaggttaacaagaatgtctata 517

RESULT 8
X60830
ID X60830 standard; DNA; 3336 BP.
XX
AC X60830;
XX
DE 09-AUG-1999 (first entry)
XX
XX S. aureus riboflavin biosynthesis operon.
XX
KM Lumazine synthase; ribH; riboflavin; differential screening; gastritis;
KM H. pylori infection; gastrointestinal carcinoma; gastric ulcer; wound;
KM bacterial adhesion; cell invasion; riboflavin biosynthesis operon; ss.
XX
OS Staphylococcus aureus.
XX
PN WO9927128-A1.
XX
PD 03-JUN-1999.
XX
PF 24-NOV-1998; 98WO-US25096.
XX
PR 25-NOV-1997; 97US-0977553.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX
```

PI Debonch C, Fedon JC, Jaworski DD, Mooney J, Palmer LM;
 PI Traini CM, Wang M, Warren RL, Zhong Y;
 XX WPI, 1999-357850/30.
 DR
 XX
 PT New *Staphylococcus aureus* ribH polynucleotide and polypeptide useful
 PT in the treatment of gastric ulcers and gastritis
 XX
 PS Example 2; Page 34-36; 48pp; English.
 XX
 CC The invention relates to polynucleotides and polypeptides of the lumazine
 CC synthase (beta -subunit) family, designated ribH. They are isolated from
 CC *Staphylococcus aureus*. ribH enzymes are involved in the synthesis of
 CC riboflavin. The polynucleotides and polypeptides may be employed as
 CC research reagents and material for the discovery of treatments and
 CC diagnostics for diseases, particularly human diseases. They can be used
 CC for diagnosis of the disease and staging of disease, and as reagents in
 CC differential screening methods. The polynucleotides may be used as a
 CC source for hybridization probes, and for screening of genetic mutations,
 CC serotype, and identification, and for chromosome identification. The
 CC polypeptides can be used to produce antibodies. The polypeptides can
 CC also be used in vaccine formulations, and to identify agonists and
 CC antagonists. These are used to prevent, inhibit or treat diseases,
 CC particularly of *Helicobacter pylori* infections, such as gastrointestinal
 CC carcinoma. Gastric ulcers and gastritis may also be treated. The
 CC polypeptides can also be used to treat wounds and in-dwelling devices to
 CC prevent bacterial adhesion and infection, and to block ribH protein-
 CC mediated mammalian cell invasion. Antagonists and agonists of ribH may
 CC also be employed to inhibit and treat diseases. The present sequence
 CC represents the *S. aureus* riboflavin biosynthesis operon that encodes
 CC genes ribG, ribB, ribA and ribH.
 XX
 XX
 SQ Sequence 3336 BP; 1202 A; 459 C; 678 G; 997 T; 0 other;

 Query Match 27.8%; Score 352.2; DB 20; Length 3336;
 Best Local Similarity 88.4%; Pred. No. 1.9e-75;
 Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;

 Oy 734 ATGATTTATGTCATTCACCACTTCACCAATATGTCACAAAGTCACAAAGGTTATTCACCC 793
 Db 1 atgattatgcatccaacttcgaacttgcacaatatgctacaagaatcaacgggttaaccacc 60

 Oy 794 GTTGCCGCTGTGTAGTTAATGTAAGTAGAGTATGTTGTTGTTGTCACACTTGAGAAAA 853
 Db 61 gtggcgctgtgttagttaatgaatgaagtagagtagtgttggtagttggcacacttggaataa 120

 Oy 854 GGTGACAAACATGCGGAGGTTCAACACTTGATATGCAACAANAATGCTGAAGTGG 913
 Db 121 ggtgacaagcatcagggaggttcaagcacttgcatalgctacaaacaatgctgaagttgcg 180

 Oy 914 ACGATTTTATTTACCTTGAAGCATTGATGTCATTTTGGTTCAACCAACCCGTGTATAC 973
 Db 181 acgatttatatagcttagagagccatgctgcaatcttggttcaaccaccccccgtgttaac 240

 Oy 974 AAAATTTATTTGATTGAAGATAGCAANAAGTAGTATTACNCAACANAAGCAATTCGTTAG 1033
 Db 241 aaaattattgattgtaagatagcaaaagtagta-tacgcaacaaaagaacatt-cgttag 298

 Oy 1034 ACACACATGGGTGATGAGACGTTACGGGGCTCCACGGTATTTGAGCGGTTGAATGCGTTG 1093
 Db 299 acacacat-ggtgatagagacgtttagggctccacggtatgaggttgaatgctt----- 351

 Oy 1094 GATGATGAAGGGCATCCCAATATACCAANAAGCTTTTTTAAACCAANAAGCAAGCAA 1153
 Db 352 gatgatgaaggggcatcccaatataccaagaactttt---aaaggcaaaagcaagcaa 408

 Oy 1154 CTTCGCACAAAATTTCACTGAAGATNTCTTGAAGTTAGATGCGTAAACAAGCGAATT 1213
 Db 409 ct--gcacaattacagtgaaagatc-tgcaggttagatggtgaaacaagcg---aat 462

 Oy 1214 GATAATGGACAAAGTCATGATTACTAACAANAAGGTTAAACAAGATGCTTATA 1268
 |||||||

Dd	463	gataatggcacaagtcacatgatctactaacaaagggttaacaagaatctcata 517						
RESULT	9							
T83787	ID	T83787 standard; DNA: 501 BP.						
XX	AC	T83787;						
XX	DT	16-JUL-1998 (first entry)						
XX	DE	DNA encoding a staphylococcus aureus protein of unknown function.						
XX	KW	Staphylococcus aureus protein; ribozyme; antisense sequence; control;						
KW	KM	staphylococcal gene; regulatory element; bacterial gene expression;						
KW	KV	vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;						
XX	KX	toxic shock syndrome; ss.						
OS	Staphylococcus aureus.							
Key	Location/Qualifiers							
FT	CDS	167..382						
FT	CDS	/+tag= a						
PN	WOJ9370070-A1.							
XX	PD	21-AUG-1997.						
XX	PF	19-FEB-1997;	97WO-US02318.					
XX	PR	20-FEB-1996;	96US-0011888.					
XX	PA	(SMIK) SMITHKLINE BEECHAM CORP.						
PI	Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO;							
PI	Pratt JM, Reichard RW, Rosenberg M, Ward JM;							
DR	WPI: 1997-424969/39.							
DR	P-PsDB: W27822.							
PT	Novel polypeptide(s) from Staphylococcus aureus strain WCUNH29 - used							
PT	to isolate antimicrobial compounds, and in vaccines against S.							
PT	aureus infection							
PS	Claim 9; Page 673; 989pp; English.							
CC	The present sequence encodes a Staphylococcus aureus protein of							
CC	unknown function. The present sequence was isolated from a							
CC	library of clones of S. aureus WCUNH 29 in Escherichia coli. The DNA							
CC	sequence can be used in the construction of ribozymes and antisense							
CC	sequences to control the expression of Staphylococcal genes. The DNA							
CC	sequence is also useful as a source of regulatory elements for the							
CC	control of bacterial gene expression. The encoded protein may be used							
CC	to produce vaccines to enable a host to produce specific antibodies							
CC	with antibacterial action. These vaccines and antibodies would protect							
CC	a host against invasion by S. aureus, and conditions relating to							
CC	Staphylococcal infection, e.g. Staphylococcal food poisoning, scaled							
CC	skin syndrome, and toxic shock syndrome.							
SQ	Sequence 501 BP; 189 A; 77 C; 94 G; 133 T; 8 other:							
OY	Query Match	18.0%;	Score 229;	DB 18;	Length 501;			
Db	Best Local Similarity	89.7%;	Pred. No. 3.9e-46;					
Matches	261; Conservative	0; Mismatches	27; Indels	3; Gaps	2;			
OY	68 CAAGGTTTGGAACATTACCAATATGATTCCTCGATGAGTCACAATGNCANCGGTGTAA	127						
I								
Db	212 ccaagtgttgcgaaccttaccacaatatgcttcgcat-aggtcaaatagtacaacgcggtgttaa	270						
OY	128 TTAACACTAGCAATAGNTGTGNAAATGATAGTAGANCAAGTGGCGCATACAGTNTCTCAATT	187						
I								
Db	271 taaactacgaatatgvtggaataatgatatqtagaacaagttcgccatcacagtcattcatt 330							


```

FT CDS 4088..4552
FT /tag= o
FT /label= B-riboflavin synthase gene
FT promoter 4545..4574
FT /tag= p
FT /label= P3 Promoter
FT RBS 4653..4659
FT /tag= q
FT CDS 4665..5039
FT /tag= r
FT /number= ORF 2
FT terminator 5038..5090
FT /tag= s
FT /standard_name= Rho-independent terminator
FT CDS complement (5567..5053)
FT /tag= t
FT /number= ORF 1
FT
XX EP405370-A.
XX
XX 02-JAN-1991.
XX
XX 22-JUN-1990; 90EP-0111916.
XX
XX 22-JUN-1989; 89US-0370378.
XX
XX (HOFF ) HOFFMANN-LA ROCHE AG.
XX
XX Perkins JB, Pero JG, Sloma A;
XX
XX WPI; 1991-008756/02.
XX
XX P-PSDB; R10569-75, R11221.
XX
XX DNA encoding riboflavin biosynthetic proteins - used to produce
XX recombinant bacteria for prodn. of riboflavin
XX
XX Disclousure; Fig 3; 72pp; English.
XX
XX The sequence or fragments thereof may be expressed in a transformed
XX bacterial expression system for the efficient production of
XX riboflavin (vitamin B2).
XX
XX Sequence 5567 BP; 1751 A; 1096 C; 1297 G; 1423 T; 0 other;
SQ

```

Query Match 8.1%; Score 103; DB 12; Length 5567;
 Best Local Similarity 60.6%; Pred. No. 1.3e-15;
 Matches 166; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

```

OY 729 AATTATGATTAATCGATTCAACTTCCAAATATGCTACAGGTGACAGGTGTTATC 788
DB 1111 attatagaagctgacctagatcttgcgaagcagcggaagacagaccgaatccaatc 1170
OY 789 CACCCTGGCGCTGTTAGTTAATGAAGTAGAGATTGTTGTTAGTGCACACTTGA 848
DB 1171 cgctgcgcgcgtctgtcgtctgaaagcagaactgtcgaatgagcgccatttaa 1230
OY 849 GAAAAGGTGACACAGCATCGGAGGTTCAAGCACTTGATGTGCACACANATGCTGAG 908
DB 1231 aataatgtaaacctcatgaaagatcattccatcattatgctggaagcaccatgcagaag 1290
OY 909 GTGCGACGATTATATAGTTAGTAGACCAATGATGATTTGTTCAACACACACCTGTG 968
DB 1291 gtgccgaaccttaacgttaacacgcgaacccgtgcagccattacggaataaacccgcattg 1350
OY 969 TTAACAAAATTAATGATTGTAAGATAGCANAACT 1002
DB 1351 cagaattgattatcaactctgtgatcaaaaagagt 1384

```

RESULT 12
 X81947
 ID X81947 standard; DNA: 5567 BP.

```

XX AC X81947;
XX 10-SEP-1999 (first entry)
XX
XX B. subtilis rib operon nucleotide sequence.
XX
XX Riboflavin; open reading frame; ORF; structural gene; promoter;
XX vitamin B2; Bacillus subtilis; rib operon; ds.
XX
XX Bacillus subtilis.
XX
XX Key Location/Qualifiers
XX CDS 1..5567
XX /tag= a
XX /note= "reading frame 1 translated protein (Y21801)"
XX CDS 2..5567
XX /tag= b
XX /note= "reading frame 2 translated protein (Y21802)"
XX CDS 3..5567
XX /tag= c
XX /note= "reading frame 3 translated protein (Y21803)"
XX
XX US925538-A.
XX
XX 20-JUL-1999.
XX
XX 24-AUG-1998; 98US-0138775.
XX
XX 11-SEP-1990; 90US-0581048.
XX
XX 22-JUN-1989; 89US-0370378.
XX
XX 21-APR-1992; 92US-0873572.
XX
XX 06-FEB-1995; 95US-0384626.
XX
XX 24-AUG-1998; 98US-0138775.
XX
XX (HOFF ) ROCHE VITAMINS INC.
XX
XX Erdemberger T, Hatch RT, Hermann T, Perkins JB;
XX Pero JG, Sloma A;
XX
XX WPI; 1999-418271/35.
XX
XX P-PSDB; Y21801, Y21802, Y21803.
XX
XX Use of bacterial strains for the over production of riboflavin
XX
XX Example 2; Fig 3A-S; 61pp; English.
XX
XX The invention relates to a method of producing riboflavin that comprises,
XX providing a recombinant bacterium containing an exogenously introduced
XX bacterial ORF (open reading frame) 3 (or ORF5) structural gene and an
XX exogenous promoter, culturing the bacterium, and recovery of the product.
XX The method is useful for the production of large quantities of riboflavin
XX (vitamin B2). The present sequence represents the complete nucleotide
XX sequence of B. subtilis rib operon (also see X81948 for the antisense
XX sequence).
XX
XX Sequence 5567 BP; 1751 A; 1096 C; 1297 G; 1423 T; 0 other;
SQ

```

Query Match 8.1%; Score 103; DB 20; Length 5567;
 Best Local Similarity 60.6%; Pred. No. 1.3e-15;
 Matches 166; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

```

OY 729 AATTATGATTAATCGATTCAACTTCCAAATATGCTACAGGTGACAGGTGTTATC 788
DB 1111 attatagaagctgacctagatcttgcgaagcagcggaagacagaccgaatccaatc 1170
OY 789 CACCCTGGCGCTGTTAGTTAATGAAGTAGAGATTGTTGTTAGTGCACACTTGA 848
DB 1171 cgctgcgcgcgtctgtcgtctgaaagcagaactgtcgaatgagcgccatttaa 1230
OY 849 GAAAAGGTGACACAGCATCGGAGGTTCAAGCACTTGATGTGCACACANATGCTGAG 908

```



```
Db 1231 aatattggtgaagctcagcagaagcttcacatccatcgtgagcagatgcagagg 1290
QY 909 GTGGGAGATTATATACGTAGACCATGTAGCATTTTGGTTACACACCCTGTG 968
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1291 gtgcgcacatttaacgttacatcgcacgcgtgcagcattacgaaaaacacgcgcagt 1350
QY 969 TTACAAATTTATGATTGTAGATAGCANAGT 1002
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1351 cagattgattatcaactctgtatcctcaagaagt 1384

RESULT 13
X81948/C
ID X81948 standard; DNA; 5567 BP.
XX
XX X81948;
XX
XX 10-SEP-1999 (first entry)
XX
XX B. subtilis rib operon complementary nucleotide sequence.
DE
XX
XX Ribo flavin; open reading frame; ORF; structural gene; promoter;
KM vitamin B2; Bacillus subtilis; rib operon; complementary; ds.
XX
XX Bacillus subtilis.
OS
XX
XX Key Location/Qualifiers
FH 1..5567
FT CDS
FT /*tag= a
FT /note= "reading frame 1 translated protein (Y21804)"
FT CDS
FT /*tag= b
FT /note= "reading frame 3 translated protein (Y21805)"
FT CDS
FT /*tag= c
FT /note= "reading frame 2 translated protein (Y21806)"
XX
XX US5925538-A.
XX
XX PD 20-JUL-1999.
XX
XX PF 24-AUG-1998; 9805-0138775.
XX
XX PR 11-SEP-1990; 9005-0581048.
XX PR 22-JUN-1989; 8905-0370378.
XX PR 21-APR-1992; 9205-0873572.
XX PR 06-FEB-1995; 9505-0384626.
XX PR 24-AUG-1998; 9805-0138775.
XX
XX PA (HOFF ) ROCHE VITAMINS INC.
XX
XX PI Eidenberger T, Hatch RT, Hermann T, Perkins JB;
PI Pero JG, Sloma A;
XX
XX DR MPI; 1999-418271/35.
XX DR P-PDB; Y21804, Y21805, Y21806.
XX
XX PT Use of bacterial strains for the over production of riboflavin
XX
XX PS Example 2; Fig 3A-S; 61pp; English.
XX
XX CC The invention relates to a method of producing riboflavin that comprises,
XX CC providing a recombinant bacterium containing an exogenously introduced
XX CC bacterial ORF (open reading frame) 3 (or ORF5) structural gene and an
XX CC exogenous promoter, culturing the bacterium, and recovery of the product.
XX CC The method is useful for the production of large quantities of riboflavin
XX CC (vitamin B2). The present sequence represents the complementary
XX CC nucleotide sequence of B. subtilis rib operon.
XX
XX SQ Sequence 5567 BP; 1422 A; 1296 C; 1097 G; 1752 T; 0 other;

Query Match 8.1%; Score 103; DB 20; Length 5567;
```

```
Best Local Similarity 60.6%; Pred. No.1.3e-15;
Matches 166; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 729 AATTATGAGTTATGCGATTCACCTCCAAATATGTCAGCAAGGCANACAGGTGTATC 788
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4457 ATTATATGAAGCTGGCCCTTGATCTTCGAAAGCGGAGGACAGACCGAATCCATC 4398
QY 789 CACCCGTTGGCGCTGTGTAGTTAATGAGTAGATGTTGTTGTTGTCACACTTGA 848
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4397 CGCTCGTGGCGCGCTGTGTGTAAAGGACGACAAATGTTCGGAATGGCCCATTTAA 4338
QY 849 GAAAAGGTGACAAAGCATCGGAGGTTCAAGCACTTGATATGCAACAAANAATGCTGAAG 908
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4337 AATATGCTGAAGCTCAGTCAGCAAGATTGATGCTATGCTGAGCAGATGCAAGG 4278
QY 909 GTGGGAGATTATATACGTAGACCATGTAGCATTTTGGTTACACACCCTGTG 968
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4277 GTGGCGACATTAGCTTACCTGACCTGACCGCTTACGGAAGAAACACCCCATGTG 4218
QY 969 TTACAAATTTATGATTGTAGATAGCANAGT 1002
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4217 CAGAATTGATTATCACTCTGTATCAAAAGAGT 4184

RESULT 14
V71600
ID V71600 standard; DNA; 5567 BP.
XX
XX AC V71600;
XX
XX DT 01-MAR-1999 (first entry)
XX
XX DE Bacillus subtilis rib operon.
XX
XX KM Ribo flavin; vitamin B2; riboflavin synthase; deaminase; reductase;
KM GTP cyclohydrolase; rib operon; ds.
XX
XX OS Bacillus subtilis strain 168.
XX
XX
XX Key Location/Qualifiers
FH 352..357
FT RBS
FT /*tag= a
FT CDS
FT /*tag= b
FT /*tag= b
FT /label= ORF6
FT terminator
FT /*tag= c
FT /note= "rho-independent transcription terminator"
FT promoter
FT /*tag= d
FT /label= PI
FT terminator
FT /*tag= e
FT /note= "rho-independent transcription terminator"
FT RBS
FT /*tag= f
FT RBS
FT /*tag= g
FT CDS
FT /*tag= h
FT /*tag= h
FT RBS
FT /*tag= i
FT CDS
FT /*tag= j
FT /label= ORF4
FT promoter
FT /*tag= k
FT RBS
FT /*tag= p2
FT /*tag= p2
FT CDS
FT /*tag= m
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 05:00:11 : Search time 111.68 Seconds
(without alignments)
1718.530 Million cell updates/sec

Title: US-08-978-456-3

Perfect score: 1269

Sequence: 1 AANCCCAATCCNATGTGGGA.....GTAAACAAGATGCTATAG 1269

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 262060 seqs, 75620727 residues

Total number of hits satisfying chosen parameters: 524120

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/1na/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/1na/5C_COMB.seq:*
4: /cgn2_6/ptodata/2/1na/5D_COMB.seq:*
5: /cgn2_6/ptodata/2/1na/5E_COMB.seq:*
6: /cgn2_6/ptodata/2/1na/5F_COMB.seq:*
7: /cgn2_6/ptodata/2/1na/5G_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1250	98.5	1269	5	US-08-978-456-3
2	352.2	27.8	1029	5	US-08-978-456-1
3	352.2	27.8	3336	3	US-08-977-554-7
4	352.2	27.8	3336	5	US-08-978-456-7
5	80.4	6.3	3417	4	US-08-978-458-7
6	80.4	6.3	3417	5	US-08-978-454-7
7	64.4	5.1	1230	4	US-08-741-327E-14
8	45.2	3.6	7218	1	US-08-232-463-14
9	41.6	3.3	19124	4	US-08-487-826B-13
10	38	3.0	731	2	US-08-451-405A-2
11	37.6	3.0	1980	4	US-08-804-794-1
12	37.2	2.9	1079	1	US-07-781-335-1
13	37.2	2.9	19124	4	US-08-487-826B-13
14	37	2.9	4248	5	US-08-678-614-1
15	37	2.9	5589	1	US-08-465-795-1
16	36.4	2.9	2317	5	US-08-749-522-5
17	36.2	2.9	2781	5	US-08-749-522-4
18	35.8	2.8	8854	5	US-09-053-549-1
19	35.6	2.8	2882	3	US-08-724-394A-12
20	35.4	2.8	665	4	US-08-883-795A-36
21	35.2	2.8	3656	4	US-08-232-463-8
22	35.2	2.8	3688	1	US-08-232-463-9
23	35.2	2.8	4659	1	US-08-232-463-10
24	35.2	2.8	4818	1	US-08-232-463-11
25	35.2	2.8	4821	1	US-08-232-463-12
26	35.2	2.8	4824	1	US-08-232-463-13

27	35.2	2.8	6601	1	US-08-232-463-15	Sequence 15, Appl
28	35.2	2.8	7218	1	US-08-232-463-14	Sequence 14, Appl
29	35.2	2.8	8710	1	US-08-480-882B-3	Sequence 3, Appl
30	35.2	2.8	8710	2	US-08-480-210-3	Sequence 3, Appl
31	35.2	2.8	9019	1	US-08-480-882B-4	Sequence 4, Appl
32	35.2	2.8	9019	2	US-08-480-210-4	Sequence 4, Appl
33	35.2	2.8	9454	1	US-08-232-463-3	Sequence 3, Appl
34	35.2	2.8	9454	1	US-08-232-463-4	Sequence 4, Appl
35	35.2	2.8	9890	1	US-08-232-463-18	Sequence 18, Appl
36	35.2	2.8	9816	1	US-08-232-463-17	Sequence 17, Appl
37	35.2	2.8	9917	1	US-08-232-463-16	Sequence 16, Appl
38	35.2	2.8	10408	1	US-08-232-463-6	Sequence 6, Appl
39	35.2	2.8	10408	1	US-08-232-463-7	Sequence 7, Appl
40	34.8	2.7	665	4	US-08-883-795A-36	Sequence 36, Appl
41	34.8	2.7	51952	5	US-08-947-823-1	Sequence 1, Appl
42	34.6	2.7	1890	7	5312912-3	Patent No. 5312912
43	34.6	2.7	3600	5	US-08-894-731-1	Sequence 1, Appl
44	34.6	2.7	51952	5	US-08-947-823-1	Sequence 1, Appl
45	34.4	2.7	1890	7	5312912-3	Patent No. 5312912

ALIGNMENTS

RESULT 1
US-08-978-456-3
Sequence 3, Application US/08978456
Patent No. 6010881
GENERAL INFORMATION:
APPLICANT: Palmer, Leslie M.
TITLE OF INVENTION: No. 6010881el ribc
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM type: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,456
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Todd Q
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50444-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2252
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1269 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-978-456-3

Query Match 98.5%: Score 1250; DB 5; Length 1269;
Best Local Similarity 100.0%: Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AANCCCAATCCNATGTGGAGGAAATCAATCAATCCCGANNCCCAATCAATCAAGTTAA 60

```
|||||
Db 1 AANCCCAATCCNATTTGGGAGSNAATCCAAATCAATNCCCGANNCCCAATCCAAAGTTAA 60
Qy 61 TTAAGTCCAGGTTTGGAAACATTAACCAATATGATTCGGATGAGTCAATGNCANCG 120
Db 61 TTAAGTCCAGGTTTGGAAACATTAACCAATATGATTCGGATGAGTCAATGNCANCG 120
Qy 121 GGTGTAATTAATACGAATGNTGTGNAAATGATAGTANANCAAGTTGGCGATACAGT 180
Db 121 GGTGTAATTAATACGAATGNTGTGNAAATGATAGTANANCAAGTTGGCGATACAGT 180
Qy 181 CTCNATTAATATGATGCTTAGAATGCAATGAGCAACAGCGAGTTCATCAATAAT 240
Db 181 CTCNATTAATATGATGCTTAGAATGCAATGAGCAACAGCGAGTTCATCAATAAT 240
Qy 241 TTGTACTAGAGATTAATAAGAGAAAGCTCTATAGAGACGAATTTGAAGTTTGA 300
Db 241 TTGTACTAGAGATTAATAAGAGAAAGCTCTATAGAGACGAATTTGAAGTTTGA 300
Qy 301 ATGTCTGTAGTAAGATATATCAATGATGATGCTATGCTAGTCAATATATTAAT 360
Db 301 ATGTCTGTAGTAAGATATATCAATGATGATGCTATGCTAGTCAATATATTAAT 360
Qy 361 AAAACCGTCATTAATTTGTTTTTGAAGAAACATATAGTATCAATTTTAATGATGACA 420
Db 361 AAAACCGTCATTAATTTGTTTTTGAAGAAACATATAGTATCAATTTTAATGATGACA 420
Qy 421 TACTAGTACTCAATATCTATTAACAATTTTCAATATATCTTTGCGGCGACGGTGAA 480
Db 421 TACTAGTACTCAATATCTATTAACAATTTTCAATATATCTTTGCGGCGACGGTGAA 480
Qy 481 ATTCCCAACCGGAGTAATAAGGCTGACCTGCTATATGTTTCAATATAGTGGCTG 540
Db 481 ATTCCCAACCGGAGTAATAAGGCTGACCTGCTATATGTTTCAATATAGTGGCTG 540
Qy 541 ATCTAGTGAATTTCTAGACCCGACAGTTAAAGTCTGATGAGGAGAAAGATTTAAT 600
Db 541 ATCTAGTGAATTTCTAGACCCGACAGTTAAAGTCTGATGAGGAGAAAGATTTAAT 600
Qy 601 CGACAAAGATTAATGATGCTATTTGTAATAATGTTGACAAATAGCTTATTTAACGATA 660
Db 601 CGACAAAGATTAATGATGCTATTTGTAATAATGTTGACAAATAGCTTATTTAACGATA 660
Qy 661 ATTTTTCCTTGCATCTTAATCTATGATGAGGATTTTGTGTTATAGAGGATCA 720
Db 661 ATTTTTCCTTGCATCTTAATCTATGATGAGGATTTTGTGTTATAGAGGATCA 720
Qy 721 TTTGAGTCAATTTATGATTAATGCGATTCACCTTCCAAATATGTTACAAAGTCAANACAG 780
Db 721 TTTGAGTCAATTTATGATTAATGCGATTCACCTTCCAAATATGTTACAAAGTCAANACAG 780
Qy 781 TGTTAATCCACCGTGGCGCTGTGTGATTAATGAGGTAGAGATTTGTTATGCTG 840
Db 781 TGTTAATCCACCGTGGCGCTGTGTGATTAATGAGGTAGAGATTTGTTATGCTG 840
Qy 841 AACCTTGACAAAAGGTGACAAAGCATGCGAGGTTCAAGACCTTGATATGCGCAACANAA 900
Db 841 AACCTTGACAAAAGGTGACAAAGCATGCGAGGTTCAAGACCTTGATATGCGCAACANAA 900
Qy 901 TCGTGAAGTGGGAGATTTATATGCTTAGAGCCATAGTACATTTGGTTCAACACC 960
Db 901 TCGTGAAGTGGGAGATTTATATGCTTAGAGCCATAGTACATTTGGTTCAACACC 960
Qy 961 ACCCTGTGTTAAACAAATTAATGATTTAGATAGCANAAAGTAGTATTCNCAACANAAG 1020
Db 961 ACCCTGTGTTAAACAAATTAATGATTTAGATAGCANAAAGTAGTATTCNCAACANAAG 1020
Qy 1021 ACAAATTCCTTAGACACACATGGGTATAGACAGTACGGGGCTCCACGGTATTTGAGGG 1080
Db 1021 ACAAATTCCTTAGACACACATGGGTATAGACAGTACGGGGCTCCACGGTATTTGAGGG 1080
Qy 1081 TTGAATGGGTGATGATGAGGACGGGACATCAATATACCAAAAGCTTTTAAAGCA 1140
|||||
```

```
Db 1081 TTGAATGGGTGATGATGAGGACGGGACATCAATTAACAAAGACTTTTAAAGCA 1140
Qy 1141 AAAAGCAAGCAACTTGGCCACAAATTAACAGTGAAGTCTTGAAGTTAGATGGGTA 1200
Db 1141 AAAAGCAAGCAACTTGGCCACAAATTAACAGTGAAGTCTTGAAGTTAGATGGGTA 1200
Qy 1201 AACAAAGCAATGATTAATGACAAAGATGATTAATCAACAAAGGTTAAACAGA 1260
Db 1201 AACAAAGCAATGATTAATGACAAAGATGATTAATCAACAAAGGTTAAACAGA 1260
Qy 1261 TGTCTATAG 1269
Db 1261 TGTCTATAG 1269

RESULT 2
US-08-978-456-1
; Sequence 1, Application us/08978456
; Patent No. 6010881
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; TITLE OF INVENTION: No. 6010881el ribG
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978.456
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd O
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50444-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1029 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-978-456-1

Query Match 27.8%; Score 352.2; DB 5; Length 1029;
Best Local Similarity 88.4%; Pred. No. 3.4e-87;
Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;
```

```
Db 121 GGTGACAAGCATGGGAGGTTCAAGCACTGTATATGCGACAACAAATGCTGAAGGTGCG 180
Qy 914 ACGATTATATTAGCTTAGACCACTGTATGCTATTTGGTCAACGCCACCTGTGTAC 973
Db 181 ACGATTATATTAGCTTAGACCACTGTATGCTATTTGGTCAACGCCACCTGTGTAC 240
Qy 974 AAAATTATGATTGATAGATAGCAANAAGTAGATTATACNCAANAGCAATTCGGTTAG 1033
Db 241 AAAATTATGATTGATAGATAGCAANAAGTAGA-TACGCACAANAAGCAATTCGTTAG 298
Qy 1034 ACACACATGGGTGATGAGACGTTACGGGGCTCCACGGTATTGAGGTTGAATGCGTTG 1093
Db 299 ACACACAT-GGTGATGAGACGTTACGGGGCTCCACGGTATTGAGGTTGAATGCGTTG 351
Qy 1094 GATGATGACGGGCGCATCATATTATACCAAGACTTTTAAAGCAAAAAGCAAGCAA 1153
Db 352 GATGATGACGGGCGCATCATATTATACCAAGACTTTT---AAAGCAAAAAGCAAGCAA 408
Qy 1154 CTGCGCACAAATTAACATGTAAGAGTCTTGAAGTTAGATGGTTAAACAAGCAAGT 1213
Db 409 CT--GCCACAATTAACATGTAAGATC-TGCAAGTTAGATGGTTAAACAAGCG---AAT 462
Qy 1214 GATAATGACAAAGTCAATGATTACTTAACAAGAAGTTAAACAAGATGCTATA 1268
Db 463 GATAATGACAAAGTCAATGATTACTTAACAAGAAGTTAAACAAGATGCTATA 517

RESULT 3
US-08-977-554-7
; Sequence 7, Application US/08977554
; Patent No. 5891672
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; APPLICANT: Fedon, Jason C.
; APPLICANT: Warren, Richard L.
; APPLICANT: Traini, Christopher M.
; APPLICANT: Wang, Min
; APPLICANT: Jaworski, Deborah D.
; APPLICANT: Mooney, Jeffrey
; APPLICANT: Debouck, Christine
; APPLICANT: Zhong, Yiyi
; APPLICANT: Black, Michael
; TITLE OF INVENTION: r1bA
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/977.554
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/02318
; FILING DATE: 19-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, O. Todd
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50444-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215/994-2252
; TELEFAX: 215/994-2222
;
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 3336 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-977-554-7

Query Match 27.8%; Score 352.2; DB 3: Length 3336;
Best Local Similarity 88.4%; Pred. No. 5.8e-87;
Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;

Qy 734 ATGATTATGCGATTCAACTTCCAAATATGTTAGTACAGGTGCAAGGTGTTAATCCACC 793
Db 1 ATGATTATGCGATTCAACTTCCAAATATGTTAGTACAGGTGCAAGGTGTTAATCCACC 60
Qy 794 GTTGGCGCTGTTAGTATTAAGAAGTATGATTTGGTATGTTGTCACACTTGAAGAAA 853
Db 61 GTTGGCGCTGTTAGTATTAAGAAGTATGATTTGGTATGTTGTCACACTTGAAGAAA 120
Qy 854 GGTGACAAGCATGGGAGGTTCAAGCACTGTATGTCACACAANAATGCTGAAGGTGCG 913
Db 121 GGTGACAAGCATGGGAGGTTCAAGCACTGTATGTCACACAANAATGCTGAAGGTGCG 180
Qy 914 ACGATTATATTAGCTTAGACCACTGTATGCTATTTGGTCAACGCCACCTGTGTAC 973
Db 181 ACGATTATATTAGCTTAGACCACTGTATGCTATTTGGTCAACGCCACCTGTGTAC 240
Qy 974 AAAATTATGATTGATAGATAGCAANAAGTAGATTATACNCAANAGCAATTCGGTTAG 1033
Db 241 AAAATTATGATTGATAGATAGCAANAAGTAGA-TACGCACAANAAGCAATTCGTTAG 298
Qy 1034 ACACACATGGGTGATGAGACGTTACGGGGCTCCACGGTATTGAGGTTGAATGCGTTG 1093
Db 299 ACACACAT-GGTGATGAGACGTTACGGGGCTCCACGGTATTGAGGTTGAATGCGTTG 351
Qy 1094 GATGATGACGGGCGCATCATATTATACCAAGACTTTTAAAGCAAAAAGCAAGCAA 1153
Db 352 GATGATGACGGGCGCATCATATTATACCAAGACTTTT---AAAGCAAAAAGCAAGCAA 408
Qy 1154 CTGCGCACAAATTAACATGTAAGAGTCTTGAAGTTAGATGGTTAAACAAGCAAGT 1213
Db 409 CT--GCCACAATTAACATGTAAGATC-TGCAAGTTAGATGGTTAAACAAGCG---AAT 462
Qy 1214 GATAATGACAAAGTCAATGATTACTTAACAAGAAGTTAAACAAGATGCTATA 1268
Db 463 GATAATGACAAAGTCAATGATTACTTAACAAGAAGTTAAACAAGATGCTATA 517

RESULT 4
US-08-978-456-7
; Sequence 7, Application US/08978456
; Patent No. 6010881
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; TITLE OF INVENTION: No. 6010881el r1bG
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978.456
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
```

```
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Todd Q
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50444-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2252
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3336 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-978-456-7

Query Match      27.8%   Score 352.2; DB 5; Length 3336;
Best Local Similarity 88.4%   Pred. No. 5.8e-87;
Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;

QY 734 ATGATTATCGATTCACTTCCAAATATGCTACAAGGTCAACAGGTGTTATCCACC 793
    |||||||
DB 1 ATGATTATCGATTCACTTCCAAATATGCTACAAGGTCAACAGGTGTTATCCACC 60

QY 794 GTTGGCGCTGTGTAGTTAATGAAGTAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 853
    |||||||
DB 61 GTTGGCGCTGTGTAGTTAATGAAGTAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 120

QY 854 GTTGACAAGATCGGGGCTTCAAGCACTGATGTCACACANAAATGCTGAAGTGGC 913
    |||||||
DB 121 GTTGACAAGATCGGGGCTTCAAGCACTGATGTCACACANAAATGCTGAAGTGGC 180

QY 914 ACGATTATATACGTAGAGCCATGATGATTTGGTTCAACACCCCTGTGTGTAC 973
    |||||||
DB 181 ACGATTATATACGTAGAGCCATGATGATTTGGTTCAACACCCCTGTGTGTAC 240

QY 974 AAAATTATTCATTGTGAAGATGACANAAAGTATTTAGTCACACANAAAGATTCGTTAG 1033
    |||||||
DB 241 AAAATTATTCATTGTGAAGATGACANAAAGTATTTAGTCACACANAAAGATTCGTTAG 298

QY 1034 ACACATATGTTGATGAGGCTTACGGGCTTCAACGCTATTTGAGGTTGAATTCGTTG 1093
    |||||||
DB 299 ACACATATGTTGATGAGGCTTACGGGCTTCAACGCTATTTGAGGTTGAATTCGTTG 351

QY 1094 GATGATGAACGGGCATCACAATTATACCAAAAGCTTTTAAAGCAAAAGCAAAAGCAA 1153
    |||||||
DB 352 GATGATGAACGGGCATCACAATTATACCAAAAGCTTTTAAAGCAAAAGCAAAAGCAA 408

QY 1154 CTGGCCACAAATATACGTAAGAGTNTCTTGAAGTTAGTGGGTAAACAAAGCAATT 1213
    |||||||
DB 409 CT--GGCACAATATACGTAAGAGTNTCTTGAAGTTAGTGGGTAAACAAAGCA--AAT 462

QY 1214 GATAAGGCAAAAGTCAATGATTTACTTAAACAAGAGGTTAAACAAGTGTCTATA 1268
    |||||||
DB 463 GATAAGGCAAAAGTCAATGATTTACTTAAACAAGAGGTTAAACAAGTGTCTATA 517

RESULT 5
US-08-978-458-7
; Sequence 7, Application US/08978458
; Patent No. 5932701
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; APPLICANT: Fedon, Jason C.
; APPLICANT: Warren, Richard L.
; APPLICANT: Kosmatka, Anna L.
; APPLICANT: Shilling, Lisa K.
; APPLICANT: Stodola, Robert K.
; APPLICANT: Knowles, David J. C.
; APPLICANT: Black, Michael T.
```

```
APPLICANT: Hodgson, John E.
APPLICANT: Nicholas, Richard O.
TITLE OF INVENTION: r1ba
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: Dechert, Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2793
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA: US/08/978,458
APPLICATION NUMBER: US/08/978,458
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/911,503
FILING DATE: 15-AUG-1997
APPLICATION NUMBER: PCT/US97/14436
FILING DATE: 15-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, O. Todd
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50533-09
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215/994-2252
TELEFAX: 215/994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3417 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-978-458-7

Query Match      6.3%   Score 80.4; DB 4; Length 3417;
Best Local Similarity 55.2%   Pred. No. 1e-12;
Matches 153; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 729 AATTATGATTTATGCGATTCACTTCCAAATATGCTACAAGGTCAACAGGTGTTATTC 788
    |||||||
DB 14 AATTATGATTTATGCGATTCACTTCCAAATATGCTACAAGGTCAACAGGTGTTATTC 73

QY 789 CACCGTGGCGCTGTGTAGTTAATGAAGTAGATGTTGTTATTTGGTGCACACTTGA 848
    |||||||
DB 74 CTTATGTTGGCGCATTTATTTAAAGATATACATTATCGACAAAGTTATTCATGAGT 133

QY 849 GAAAGGTGACACAGATCGGAGGTTCAAGCACTTGATATGSCACAACANAAATGCTGAAG 908
    |||||||
DB 134 TTTTGTGGGCCACATGCTGAGAGAAATGCTTAAACATGTAAGAAATCCCTGTGCG 193

QY 909 GTGCGACGATTTATATACGTTAGAGCCATGTAGTATTTTGGTTCAACACACCTGTG 968
    |||||||
DB 194 GAGCGAGCGCTTATATGAACCTTGAACCTGTTGTCACCTGGGAAACACACCTCCCTGTA 253

QY 969 TTACAAATTTATGTTAGTATGATGACANAAAGTAGT 1005
    |||||||
DB 254 TAGATGCTATATGATGATGATGATTTACAAAGAGTAGT 290

RESULT 6
US-08-978-454-7
; Sequence 7, Application US/08978454
; Patent No. 6017728
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
```



```
APPLICANT: Fedon, Jason C.
APPLICANT: Warren, Richard L.
APPLICANT: Kosmatka, Anna L.
APPLICANT: Shilling, Lisa K.
APPLICANT: Stodola, Robert K.
APPLICANT: Knowles, David J. C.
APPLICANT: Black, Michael T.
APPLICANT: Hodgson, John E.
APPLICANT: Nicholas, Richard O.
TITLE OF INVENTION: r1dh
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: Dechert, Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2793
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,454
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/911,503
FILING DATE: 15-AUG-1997
APPLICATION NUMBER: PCT/US97/14436
FILING DATE: 15-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Q. Todd
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50533-11
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215/994-2252
TELEFAX: 215/994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3417 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-978-454-7

Query Match          6.3%; Score 80.4; DB 5: Length 3417;
Best Local Similarity 55.2%; Pred. No. 1e-12; Indels 0; Gaps 0;
Matches 153; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 729 AATTATGATTAATGCGATCACTCCAAATATGTACAGGTCANACGGTGTATATC 788
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 14 AATATATGAATTAAGCAATTAATGCGACAAAAGGGCTGTTCGTCATATCCCAATC 73
QY 789 CACCCGTTGGCCCTGTGTAGTTAATGAAGTAGAGATTGTGTATGTGTCACACTTGA 848
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 74 CTAATGTTGGCGCAATTAATGTTAAAGATAATCAATTTATCGACAAGTTATCATGAGT 133
QY 849 GAAAGGTACACAAGCATGCGGAGGTCACAGCACTTATATGCAACAANAATGCTGAAG 908
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 134 TTTTGTGGTGGCCACATGCTGAGAGAAATGCTTTAAAAAACGTGTAGAAAATCCCTGTG 193
QY 909 GTGCGAGATTATATATAGTTAGAGCCATGTATGATTTTGTGAACACCAACCTGTG 968
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 194 GAGCGAGGCTTTATGTAACACTGAACCTGTTGTCACTTCGGGAAAACACCTCCTGTA 253
QY 969 TTAACAAAATTTATGATTTAGATAGACANAAGTACT 1005
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 254 TAGATGCTATATGATAGTATGATATTAACAAGTACT 290
```

```
RESULT 7
US-08-741-327E-14
Sequence 14, Application US/08741327E
Patent No. 5925354
GENERAL INFORMATION:
APPLICANT: Fuller, et al.
TITLE OF INVENTION: against Actinobacillus pleuropneumoniae
TITLE OF INVENTION: against Actinobacillus pleuropneumoniae
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESS: G. Kenneth Smith
STREET: 300 S. Wacker Drive
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/741,327E
FILING DATE: October 28, 1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: G. Kenneth Smith
REGISTRATION NUMBER: 43,135
REFERENCE/DOCKET NUMBER: 97704-E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-9808
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1230 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-741-327E-14

Query Match          5.1%; Score 64.4; DB 4: Length 1230;
Best Local Similarity 51.6%; Pred. No. 1.5e-08;
Matches 143; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 729 AATTATGATTAATGCGATCACTCCAAATATGTGTACAGGTCANACGGTGTATATC 788
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 152 ACTATATGCGCGCGCATTCAGCTGGCAAAACAAGTTAGCTGGAGCAATGCCAATC 211
QY 789 CACCCGTTGGCCCTGTGTAGTTAATGAAGTAGAGATTGTGTATGTGTCACACTTGA 848
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 212 CCGTTGTCGGTGTGAATTTGCAAAAACGGTGAATCGTCCGAAAGTTACCATGAAA 271
QY 849 GAAAGGTACACAAGCATGCGGAGGTTCAAGCACTTATATGACACAANAATGCTGAAG 908
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 272 AGATTGTGTAGATGCGAGCAACGTAATGCCGTTTATCATTTGTAAGAAATCTTTCCG 331
QY 909 GTGCGAGATTATATATAGTTAGAGCCATGTATGATTTTGTTCACACCAACCTGTG 968
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 332 GGGCGAGCTCTTATGTAACGCTTGAAGCCTTGTGTGATCAGCGCGACGCGCCTGTT 391
QY 969 TTAACAAAATTTATGATTTAGATAGCANAAAGTACT 1005
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 392 CGATTATTAATTTGAACGAGCATTAATAAAGATAT 428

RESULT 8
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
```

APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZ9PL-F15
US-08-232-463-14

Query Match 3.6%; Score 45.2; DB 1; Length 7218;
Best Local Similarity 3.6%; Pred. No. 0.006; Mismatches 164; Indels 0; Gaps 0;
Matches 14; Conservative 211;
QY 869 GAGGTTCAAGCCTGATGATGACACACANAAATGCTGAAGGTGGAGCATTTAATATTACG 928
DB 1454 GAGATAGAGAAATTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1395
QY 929 TTAAGCCATAGTACATTTGGTTCAACACACCCTGCTGTAACAATAATATTGATGT 988
DB 1394 RRR 1335
QY 989 AAGATAGCANAAAGTAGTTTCAACANAAAGACAATCCGTTAGACACATGGGAT 1048
DB 1334 RRR 1275
QY 1049 GAGACGTACGGGCTCCACGTAATTTGAGGGTGAATTTGGTTGGATGATGAAGCGCA 1108
DB 1274 RRR 1215
QY 1109 TCACAATTATCAACAAGACTTTTAAAGCAAAAGCAAGCAACTGCCACAAATTA 1168
DB 1214 RRR 1155
QY 1169 CAGTGAAGTTCCTGAAGATTAGATGGTAACAAGCAATGATAATGACCAAGT 1228
DB 1154 RRR 1095
QY 1229 CAATGATTACTAACAAGAGTTAAACA 1257

DB 1094 RRR 1066
RESULT 9
US-08-487-826B-13
Sequence 13, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knodbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match 3.3%; Score 41.6; DB 4; Length 19124;
Best Local Similarity 51.0%; Pred. No. 0.09; Mismatches 94; Indels 0; Gaps 0;
Matches 98; Conservative 0;
QY 282 AATGAGGTTGATTTTAAATGCTGTAGTAAGAAATCATATCAATGAGATGCTAATG 341
DB 2072 AAAATTAATTTGTTTATTAATTAATGACTTAATAATATTGTTGATTAACATATGTAAT 2131
QY 342 ACACGATTAATTAATTAATTAACCGTCATTAATTTGTTTAAAGAAACATATAGTATC 401
DB 2132 TCATTTAACAGAAATAAATAAATATATATATATATATATTAATTAATTAAGTTAT 2191
QY 402 ATTTAAATGATGACATACATGACTGACATCAATTAATTAATTAATTAATTAATTAAT 461
DB 2192 AATAAATAATATATACATATGATTAATAAATGAAGTTCACTACGATGAATATATAT 2251
QY 462 TCTTTGGGGCA 473
DB 2252 TATATGTGTCA 2263

ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/781.355
FILING DATE: 19911022
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: UTSH:162
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1079 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: linear
US-07-781-355-1

Query Match 2.9%; Score 37.2; DB 1; Length 1079;
Best Local Similarity 50.0%; Pred. No. 0.39;
Matches 93; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

OY 283 AATGAGGTTGATTTAATGCTGTTAGTAAGAAATCATATGAGATGCGCTAAT 341
DB 277 AACAGTAGATTTTATTTATTTTTCATATTTATCTCTTAATATATGCTTAAAT 218
OY 342 ACTCAGATATATTTAATTAACCGCTCATTAATGTTTTTTAGAAACATATAGTAC 401
DB 217 TATATTTTATATCAATTTATTTTCAATATCAATTTAATTAATCAATATTTTA 158
OY 402 ATTTAATAGTGTGACATGCTACTCAATATATCTATATACATTTCAATTAAT 461
DB 157 AATCATATATTTTAAATATCATATATTTTAAATATCATATTTTAAATATCAT 98
OY 462 TCTTTC 467
DB 97 GCCTAC 92

RESULT 13
US-08-487-826B-13/C
Sequence 13, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitais, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match 2.9%; Score 37.2; DB 4; Length 19124;
Best Local Similarity 47.3%; Pred. No. 1.4;
Matches 134; Conservative 0; Mismatches 148; Indels 1; Gaps 1;

OY 123 GTTATTAACATCGAATGNTGTGNAATGATAGTACGAACTGGCCATACAGTNTCT 182
DB 1520 GTTACTAATATTAATATATTTTATGCAAAATGTTTGAAGTTTAAAGATTAT 1461
OY 183 CNATTATATGATGCTTTGAATCGAATGAGCAACACAGCCAGTTACATCAATATTT 242
DB 1460 TATTTTAAATCAATATATATATGAAATAATATTAATATATATATATAT 1401
OY 243 GTAATAGAGTAAATTAAGAGAACGCTCTATAGAGAGATTTGAGTTGATTTAAT 302
DB 1400 TTAATAAAAAAAAAAGAGAGAACATTTAATAAAGATTAATATATATATATAT 1341
OY 303 GTCGTAGTAAGATCATATCAATGAGATGCTATAGTACTGATATATTAATTA 362
DB 1340 AGATACCGATTTGATATTTCTTTAGTCTGTATGTATATACA-AATATATTTATTA 1282
OY 363 AACCGTATTAATGTTTTTTTGAATAACATATAGTATCATTT 405
DB 1281 TATTTACATATTAATATATTTTGTATACATATCTAATATAT 1239

RESULT 14
US-08-678-614-1
Sequence 14, Application US/08678614
Patent No. 6013507
GENERAL INFORMATION:
APPLICANT: Tomasz, Alexander
APPLICANT: Delencastre, Hermia
TITLE OF INVENTION: AUXILIARY GENES AND PROTEINS OF
TITLE OF INVENTION: METHICILLIN RESISTANT BACTERIA AND ANTAGONISTS THEREOF
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
STREET: floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/678,614
FILING DATE: 10-JUL-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-139
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4248 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Staphylococcus aureus
STRAIN: RUSA 266
FEATURE:
NAME/KEY: ORF
LOCATION: 154..1410
NAME/KEY: ORF
LOCATION: 1497..3500
US-08-678-614-1
```

Query Match 2.9% Score 37; DB 5; Length 4248;

Best Local Similarity 52.3%; Pred. No. 0.82;

Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

```
QY 757 AATATGTCACAGTCACAGTGTATATCCACCCGCTGTTAGTAAAGA 816
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1943 AACTAAATTAACACCTGATTAATGATTAATCAAGCACAATGCTGATTAATACTAA 2002
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 817 AGTAGAGTTTGTGTAATGTCACACTTGAGAAAGGTGACAAGCATGGCGAGTTCA 876
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2003 AGCTCAAGAAAGGTTACAGACGACACAAAGTAAAGTTACGACAGATGAAGTTGC 2062
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 877 AGCACTGATATGCGACACACANAATGCTGAA 907
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2063 AGCACTGTACTAAATTAACATGATTA 2093
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 15

US-08-465-795-1

; Sequence 1, Application US/08465795

; Patent No. 5589355

; GENERAL INFORMATION:

; APPLICANT: Koizumi, Satoshi

; APPLICANT: Yonetani, Yoshiyuki

; APPLICANT: Teshiba, Sadao

; TITLE OF INVENTION: A Process for Producing Riboflavin

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Finnegan, Henderson, Farbow, Garrett &

; ADDRESSEE: Dunner

; STREET: 1300 I Street, N.W., Suite 700

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20005-3315

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,795
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER: US 08/161,394
FILING DATE: 06-DEC-1993
APPLICATION NUMBER: JP 326578/1992
FILING DATE: 07-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Gaybrick, Robert J.
REGISTRATION NUMBER: 27,890
REFERENCE/DOCKET NUMBER: 04853.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5589 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Corynebacterium ammoniagenes
US-08-465-795-1
```

Query Match 2.9% Score 37; DB 1; Length 5589;

Best Local Similarity 49.2%; Pred. No. 0.93;

Matches 122; Conservative 0; Mismatches 123; Indels 3; Gaps 1;

```
QY 764 GTACAGGTCAACAGGTGTAATCAACCCGCTGCTGTAGT---AATGAAGT 820
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 672 GTCCGCGGACTACACCCGGAATCCACCGGTGGTCCGTAATTAATGACTCCGGT 731
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 821 AGATTTGTTGTTATGTCACACTTGAGAAAGGTGACAAAGCATGGCGAGTTCAAGCA 880
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 732 GAGATTTGTGGCACCGCTGCGACTACCGCGTGGCGGGGTGCGACGAGAATCCAAAGCT 791
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 881 CTTATATGCGACACAAANATGCTGAGGTGCGACGATTTATATAGTAGAGCCATGT 940
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 792 CTAGCCGATGCGCGGCGGCAAGCCGAGCGCTACCGCGTGTGACGCTGGAGCCGCTGC 851
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 941 AGTCAATTTGTTCAACACCCCTGTGTTACAAATTAATTAAGATAGCANAA 1000
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 852 CGGATACCGCGCCGACGAGCGCTGCACGACGCTTAATTAAGCCGCGCATCAAGAT 911
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1001 GTAGTATT 1008
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 912 GTCCTTTT 919
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: November 19, 2000, 05:00:48

Job time: 8885 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 03:36:00 ; Search time 1232.02 Seconds
(without alignments)
6368.388 Million cell updates/sec

Title: US-08-978-456-3

Perfect score: 1269
Sequence: 1 AANCAACCATTCNMTTGGGA.....GTTAACAAGATGCTATAG 1269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7189864 seqs, 3091403243 residues

Total number of hits satisfying chosen parameters: 14379728

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: em_est40:*
42: em_estfun:*
43: em_esthum1:*

44: em_esthum2:*
45: em_esthum3:*
46: em_esthum4:*
47: em_esthum5:*
48: em_esthum6:*
49: em_esthum7:*
50: em_esthum8:*
51: em_esthum9:*
52: em_esthum10:*
53: em_esthum11:*
54: em_esthum12:*
55: em_esthum13:*
56: em_esthum14:*
57: em_esthum15:*
58: em_esthum16:*
59: em_esthum17:*
60: em_esthum18:*
61: em_esthum19:*
62: em_esthum20:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estom:*
68: em_estov1:*
69: em_estrov2:*
70: em_estp1:*
71: em_estcp12:*
72: em_estp13:*
73: em_estp14:*
74: em_estp15:*
75: em_estro1:*
76: em_estro2:*
77: em_estro3:*
78: em_estro4:*
79: em_estro5:*
80: em_estro6:*
81: em_estro7:*
82: em_estro8:*
83: em_estro9:*
84: em_estro10:*
85: em_estro11:*
86: em_estro12:*
87: em_estro13:*
88: gb_gss1:*
89: gb_gss2:*
90: gb_gss3:*
91: gb_gss4:*
92: em_gss1:*
93: em_gss2:*
94: em_gss3:*
95: em_gss4:*
96: gb_gss5:*
97: gb_gss6:*
98: gb_gss7:*
99: gb_gss8:*
100: gb_gss9:*
101: em_gss5:*
102: em_gss6:*
103: em_gss7:*
104: em_gss8:*
105: em_gss9:*
106: em_gss10:*
107: em_gss11:*
108: gb_gss10:*
109: gb_gss11:*
110: em_gss12:*
111: gb_gss12:*
112: gb_gss13:*
113: gb_gss14:*
114: gb_gss15:*
115: gb_gss16:*
116: gb_gss17:*

LOCUS	CNE030396	1101 bp	DNA	GSS	03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR08K10 of RpCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL063921				
VERSION	AL063921.1 GI:4941778				
KEYWORDS	GSS.				
SOURCE	fruit fly.				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;				

REFERENCE
AUTHORS
TITLE
Direct Submission
1 (bases 1 to 1101)
Genscope.
1000
[Genscope National] 20
Genscope

Query Match	Best Local Similarity	19.7%	Pred. No. 1,1e-06;	Matches 140;	Conservative 276;	Mismatches 220;	Indels 4;	Gaps 2;
QY 88	AAATGATGTCGGATGAGTCGCAATGNCANCGGTGTTAATAACTACGAATGTTGTCN	147						
Db 1098	RRMGDDTMDRDTBKDKDDMTKMTTWMKRRADRRWAGDADWMDGATGTTATWMMW	1039						
QY 148	AAATGATGTAGANCAAGTTGCCATACAGNCTCMTATATGATGCTTTAGATGCA	207						
Db 1038	WMMATWDTWMDKMMWMAATATDTAATWTRKRAMRDWGRGAGKRPDRDATDIDGAGR	979						
QY 208	ATGAGCAACACAGCGCTTACATCAATCAATATTTGTACTAGAGAATAAAGAGAC	267						
Db 978	RDGGRKRDKDRKDDGDDKRGKKKRAKAKAMATKMWDMDDMDKMKWGDANDRKADD	919						
QY 268	GCTCTATAGAGACGAATTGAAGTTTGAATTTTAATGTCTGTAC-TAAGAATCATATCAA	326						
Db 918	DDGAGDKDDDKGKADDDDTGTGTDKDDDKDKWDDMDKAGTGWGDATWMAATIDWMMWGA	859						
QY 327	TGAGATGCTTACTAGTACGACATATATATTAATTAACCGTCATTAATGTTTTTTTGA	386						

[illegible][illegible]

RESULT	3	
CNS0155H	1001 bp	DNA
LOCUS		GSS
DEFINITION	Drosophila melanogaster genome survey sequence SP6 end of BAC BACN13C23 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.	26-JUL-1999
ACCESSION	AL105023	
VERSION	AL105023.1	GI:5617037
KEYWORDS	GSS.	
SOURCE	fruit fly.	
ORGANISM	Drosophila melanogaster	
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
AUTHORS	1 (bases 1 to 1001)	
TITLE	Genoscope.	
JOURNAL	Direct Submission	
COMMENT	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage; BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr - Web : www.genoscope.cns.fr)	
FEATURES	Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CPBH (Centre d'etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.	
source	1..1001	Location/Qualifiers
	/organism="Drosophila melanogaster"	
	/plasmid="pBelobAC11"	
	/db_xref="taxon:7227"	
	/clone_11b="DrosBAC"	
	/clone="BACN13C23"	
	/note="end : SP6"	
BASE COUNT	266 a 219 c 134 g 150 t	232 others
ORIGIN		
Query Match	4.7%; Score 60.2; DB 121; Length 1001;	
Best Local Similarity	34.4%; Pred. No. 0.0013;	
Matches 117; Conservative 61; Mismatches 162; Indels 0; Gaps 0.		
109	AAATGNCANCGGTCTTAATACTACGAAATGNTGTAATGATAGTACGACGAAGTTG	168
111	: : : : : : : : : : : :	
Db	AAANVHNMAWMAAARAKAARRRAAAAAAAGDARAAARAAAGRAAARARAAAAAATTA	715
169	CGCATACAGTCTCCTNATTAATGATGCTTGAATTCGATGAGCAACACAGCGCAGTT	228
171	: : : : : : : : : :	
Db	AAWAAATAARWAAARADARAAAAAATTAATTAAMWAAWAAWMAAAATTAADTTWTAT	775
229	ACAATCAATTAATTTGTAAGTACGACATATAAAGAAGACGCTCTATAGACGAATGAA	288
231	: : : : : : : : : :	
Db	WAAATTAATAAAAAAAMWTAATAATTTAAWMAWMAAAAAAATWMAAAAMWTAMWA	835
269	GGTTGATTTTATCTCGTCTAGTAAAGATCATCAATGACATGCGCTTACTAGCCGA	348
271	: : : : : : : : : :	
Db	AAAAAAATTTAAATTTTAAWMAWMAWMAWTAATATAAADAADAAAAAAMAAWATA	895
349	TTATATTTAAATTAACCGTCATTTATGTTTTTTTGAAGAAACATATGATCATTTTAA	408
351	: : : : : : : : : :	
Db	AAATTTAAAAATTTTWTAMWADMTWTTTTWTWTWMAAAMAAATTAAMTWTATTAATWTA	955

Y	409	ATGTAGTGCATCTACTACTCAATCAATATCTTATACCA	448
		: : : :	
Db	956	WATATAAATATAAAAAATWAAWAAWATATWATATAAAAAA	995
RESULT	4		
CNS00E07			
LOCUS			
DEFINITION		CNS00E07 1101 bp DNA GSS 04-JUN-1999	
		Drosophila melanogaster genome survey sequence TET3 and of BAC:	
		BACR29P01 of RPCL-98 library from Drosophila melanogaster (fruit	
		fly), genomic survey sequence.	
ACCESSION		AL069440	
VERSION		AL069440.1 GI:4949583	
KEYWORDS		GSS.	
SOURCE		fruit fly.	
ORGANISM		Drosophila melanogaster	
		Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
		Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;	
		Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
REFERENCE		1 (bases 1 to 1101)	
AUTHORS		Genoscope.	
TITLE		Direct Submission	
JOURNAL		Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage	
		Bp 191 91006 EVRI cedex - FRANCE (E-mail : segre@genoscope.cns.fr	
COMMENT		- Web : www.genoscope.cns.fr)	
		Determination of this BAC-end sequence was carried out as part of a	
		collaboration with the Berkeley Drosophila Genome Project (BDGP).	
		The BDGP is constructing a physical map of the Drosophila	
		melanogaster genome using these BACs. For further information	
		please see http://www.fruitfly.org The BDGP Drosophila	
		melanogaster BAC library was prepared by Kazutoyo Osoegawa and	
		Aaron Mammoss in Pieter de Jong's laboratory in the Department of	
		Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,	
		NY. The library is named RPCL-98 and was constructed by partial	
		ECORI digestion of Drosophila DNA provided by the BDGP from the	
		isogenic strain y2: cn bw sp, the same strain used for the BDGP's	
		p1 and EST libraries. A more detailed description of the library	
		and how to order individual BAC clones, the entire library, or	
		filters for hybridization from the BACPAC Resource Center can be	
		found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .	
FEATURES		Location/Qualifiers	
source		1..1101	
		/organism="Drosophila melanogaster"	
		/db_xref="taxon:7227"	
		/clone_lib="RPCL-98"	
		/clone="BACR29P01"	
		/note="end : TET3"	
BASE COUNT		366 a 66 c 104 g 351 t 214 others	
ORIGIN			
Query Match		4.7% ; Score 59.8 ; DB 121 ; Length 1101 ;	
Best Local Similarity		35.1% ; Pred. No. 0.0016 ;	
Matches		202 ; Conservative 77 ; Mismatches 296 ; Indels 1 ; Gaps 1 ;	
Y	190	ATGATGCTTTAGATCGANTGAGCAACAACAGCGAGTTACATCAATATTTGTAAC	249
		: : :	
Db	442	AKAAATATTTKGAAGATATAAAAAAARATAAWTAATATATAAATWATATAA	501
Y	250	GAAGATATAAGACAAGCGCTCTATAGAGACATGAAGCTTGATTTTAAATGCTGT	309
		: :	
Db	502	WTTTAAAAAAMAAATAAATAAATCAAAATWTAATAATATTTTAAATAAMWMAATKT	561
Y	310	AGTAAGAATCATATCATGAGATGCCCTATGACTCAGATTTATTAATTAACCGCTG	369
		: : : :	
Db	562	TTTATWMAATATAAAAAAATTTATTTTATTTTATTTAAATWMAAGWMAAAAAAAMT	621
Y	370	ATTATTTGTTTTTATGAAAAATATAGATCATTTTAAATGTAAGTGCATAC	429
		: : : : :	
Db	632	ATAAHTWTAATTTATATWTTWMAATTTTAAATTAATTAATTTAAWTTWTTTAA	681
Y	430	CTCAATATATCATATACATTTCAATATATATCTTTCGGCGCAGGCTGAATTC	489

Db	682	WTTTAAMFATTAARWMAATTA	AAAAAAMWTTTAAATAA	MAAATAAATAATTTT	741
Qy	490	CGCGAGAAATAAAGCG	CGACCGCAATATATTCAT	TATAGTCGATAGTGA	549
Db	742	TTTATTTAAWMAATTA	AAAAATTWTAATMAAMWTT	MAAAAAATNTATATAA	801
Qy	550	GATTCAGACCGACAGT	TAAGTCGGATGGGAAAGA	ATGTTAAATTCACAA	609
Db	802	-WKAATAATWATAMW	TTTAAATWATATTAAT	TAATWTTTTTWTATTA	860
Qy	610	TAAATGAGCGTATTTG	AAAAATGTGTCAATAGCG	CTAATTAACATTAAT	669
Db	861	TTTTTTTAAWTAAMW	GTAAATMAAATTAAMW	TTTAAAAAATGTTT	920
Qy	670	CTTGCATCTTAATTC	ATCATGATGTGAGAT	TTTTTGTATTAAGCG	729
Db	921	WTTTTATTTWMAAT	TGTATTAATAAAAAA	MAATATGATATMAA	980
Qy	730	ATTATGAGTATAGCAT	TCACACTCCAAATATG	765	
Db	981	AATATGCRATATTT	AAAAAATAATWTA	AAAAAMAT	1016

	RESULT	5	CNS00396/c	DNA	GSS	03-JUN-1999
	LOCUS		CNS00396	1101 bp		
	DEFINITION		Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR08M12 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.			
	ACCESSION		AL063931			
	VERSION		AL063931.1	GI:4941788		
	KEYWORDS		GSS.			
	SOURCE		Fruit fly.			
	ORGANISM		Drosophila melanogaster			
			Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
			Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;			
			Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
	REFERENCE		1 (bases 1 to 1101)			
	AUTHORS		Genoscope.			
	TITLE		Direct Submission			
	JOURNAL		Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :			
			BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr			
			- Web : www.genoscope.cns.fr)			
	COMMENT		Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACS. For further information please see http://www.fruityfly.org/The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Oosegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp. the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.			
	FEATURES		Location/Qualifiers			
	SOURCE		1..1101			
			/organism="Drosophila melanogaster"			
			/db_xref="taxon:7227"			
			/clone_lid="RPCI-98"			
			/clone="BACR08M12"			
			/note="end : TET3"			
	BASE COUNT		260 a	155 c	97 g	385 t 204 others
	ORIGIN					
	Query Match		4.6%;	Score 59:	DB 121;	Length 1101;
	Best Local Similarity		37.6%;	Pred. No. 0.0024;		
	Matches 120;		Conservative 49;	Mismatches 150;	Indels 0;	Gaps 0;

Query Match	4.6%	Score 57.8	DB 113	Length 337
Best Local Similarity	54.6%	Pred. No. 0.0042		
LOCUS	AO991504	337 bp	DNA	GSS
DEFINITION	RFc02461	Photorhabdus luminescens strain W14 M13 library		
ACCESSION	AO991504	Photorhabdus luminescens genomic clone PLG02461, DNA sequence.		
VERSION	AO991504.1	GI:9650098		
KEYWORDS	GSS.			
SOURCE	Photorhabdus luminescens.			
ORGANISM	Photorhabdus luminescens			
REFERENCE	1 (bases 1 to 337)			
AUTHORS	ifrench-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T., Daborn,P., Bowen,D. and Blattner,F.R.			
TITLE	A genomic sample sequence of the entomopathogenic bacterium Photorhabdus luminescens W14: potential implications for virulence			
JOURNAL	Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)			
COMMENT	Contact: ifrench-Constant RH Department of Biology and Biochemistry University of Bath South Building, Bath BA2 7AY, UK Tel: (44) 1225 826621 Fax: (44) 1225 826779 Email: bssr@icbath.ac.uk This is one of 2,122 random reads from the M13 library. For annotation of identified clones (BLASTX, BLASTN and mapping to E. coli K12 genome) please see ifrench-Constant et al. 2000, Nucleic Acids Res. Seq primer: M13 Forward Class: Shotgun.			
FEATURES	Location/Qualifiers			
source	1..337			
	/organism="Photorhabdus luminescens"			
	/strain="W14"			
	/db_xref="taxon:29488"			
	/clone="PLG02461"			
	/clone_lib="Photorhabdus luminescens strain W14 M13 library"			
	/dev_stage="primary phase variant"			
	/note="genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus."			
BASE COUNT	86 a 60 c 93 g 97 t 1 others			
ORIGIN				

Matches	113:	Conservative	0:	Mismatches	94:	Indels	0:	Gaps	0:
QY	729	AATTATGATTAATGCGATTCACCTTCCAAATATGCTCAAGGTCANACAGGTGTTATC							788
Db	125	AATATATGCGCCGCTCTTTGGAAATTAGCGCTCAAGGCGCTTACACGTCACCTATTC							184
QY	789	CACCGGTGGCGGCTGTTAGTATGAAGGATGTTGGATTGTCACACTGA							848
Db	185	CCACGTATGCTTGTGTATGTTTGCCTGATGGCAGATTAATTGGGAAAGTTTCTATTAC							244
QY	849	GAAGAAGTACAGCATCGAGGCTTCAAGCATCTGATATGCAACAMACATGCTGAAG							908
Db	245	GTTCTGGCGGACCCACATCGGAGGTTCAATGCTTACGAGTGGCGGAGATAAG							304
QY	909	GTCGACGATTTATTAATTAAGTTAGAC							935
Db	305	GCGCAACTGTTTATGTCACGCTTGACC							331
RESULT	7								
LOCUS	CNS00LT2								
DEFINITION	CNS00LT2 1101 bp DNA GSS 14-JUN-1999								
ACCESSION	Drosophila melanogaster genome survey sequence TERT end of BAC: BACR48P19 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.								
VERSION	AL078714.1 GI:5102004								
KEYWORDS	GSS.								
SOURCE	fruit fly.								
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.								
REFERENCE	1 (bases 1 to 1101) Genoscope.								
AUTHORS	Direct Submission								
TITLE	Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage								
JOURNAL	Bp 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr								
COMMENT	- Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoler in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .								
FEATURES	location/Qualifiers								
SOURCE	1..1101								
	/organism="Drosophila melanogaster"								
	/db_xref="taxon:7227"								
	/clone_11b="RPCI-98"								
	/clone="BACR48P19"								
	/note="end : TERT"								
BASE COUNT	469 a 6 c 69 g 151 t 406 others								
ORIGIN									
Query Match	4.68; Score 57.8; DB 121; Length 1101;								
Best Local Similarity	24.48; Pred. No. 0.0044;								
Matches 158; Conservative 188; Mismatches 256; Indels 6; Gaps 2;									
QY	200	AGAATCGAATAGCACACAGCGCAGTTACATCAATTAATTTTGTACTAGAGAATAATA							259
Db	400	AA							459

```

QY 260 AAGAGACGCTCTATAGACGCAATTGAAGTTGATTTTAATGCTCTAGTACGAAATC 319
    || || || || || || || || || || || || || || || || || || || || ||
Db 460 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 519
QY 320 ATATCAATGAGATGCTATAGTACAGATTAATTAATTAATTAATTAATTAATTAAT 379
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 520 WWDTTTTTAAWTTTWTWTTAAATATTTTTTTTTTTTWTWTKAKTWTWTKAKTAATAA 579
QY 380 TTTTGAAGAAACATATAGTATCATTTTAAATAGTACGATCTACGACTCAATAAT 439
    || || || || || || || || || || || || || || || || || || || || ||
Db 580 TTTTAAATWTTTWTWTTAKTMAKMAKMAKMAKMAKMAKMAKMAKMAKMAKMAKMA 639
QY 440 CTATACATTTTCATATATATTTCTCGGGCAGGCGAATCCCAACCGCAGTAA 499
    | : : || || | : || || | : || || | : || || | : || || | : ||
Db 640 WTTTAAWTTTTTTTTTTAKTKTTTTTTTTAA--TTAAATAAAAAAATTTDTPMAAM 696
QY 500 TTAACCTCGCAGCTCTATATGTTTCATATAGTACGATCTAGTACGATCTAGAG 559
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 697 WTTTTRKKKKKAAADKWDAAKRWDAKRAKTKTKKDKKMAAADAADKDKKGGKKKK 756
QY 560 CCGACAGTTAAAGTCGATGGAGAGAAAGATTTAATATGACAAAGATATAGTACG 619
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 757 KKKGGGKKKKKKKKKKGGGKKKKKK---KAGDDAKDKTKKKKKKAATTTTCKKKKK 813
QY 620 TATTTGTAATAATGTACAAATAGCTTATTTAAGATMAATTTTCTCTTGATCT 679
    : : : || || : : || : : || : : || : : || : : || : : || : : ||
Db 814 KKAARKKKAADRTKTKMDAAAAAAAKTKDKKKKKKKTKKKKKKKKKKKKKKKKKKK 873
QY 680 TAATTCATGATGTGAGATTTTGTGTTATAGAGTGATCATTTGAGTCAATTTATGAT 739
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 874 DAAAKKKKKGTKKKKKKKKGGGKKGGKKGGKDDDAAAAKKKKKKKGGGKKKKKK 933
QY 740 TATGGATTCACATCCAAATATGTAGTACAGTTCANACAGTGTATATCCACCGTGGC 799
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 934 GKKKKGGGKKKKTKTKKKKKKKKKKADAKKTKKKRAAADAADAADWDIAKTKKKKK 993
QY 800 GCTGTGTAGTAAATGAAGTAGATTTGTTGATTTGTCACACTTG 847
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 994 DKKKTKTKKKKDAATAATKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 1041

RESULT 8
CNS00FEA/c 1101 bp DNA GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC:
DEFINITION BACR31A09 of RPCL-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION AL070314.1 GI:4950558
VERSION AL070314.1
KEYWORDS fruit fly.
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
          Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
          Muscomorpha; Ephyridioidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
          BP 191 91006 Evry cedex - FRANCE (E-mail : sequefgenoscope.cns.fr
          - Web : www.genoscope.cns.fr)
          Determination of this BAC-end sequence was carried out as part of a
          collaboration with the Berkeley Drosophila Genome Project (BDGP).
          The BDGP is constructing a physical map of the Drosophila
          melanogaster genome using these BACs. For further information
          please see http://www.fruitfly.org The BDGP Drosophila
          melanogaster BAC library was prepared by Kazuhiro Osoegawa and
          Aaron Mamoser in Pieter de Jong's laboratory in the Department of
          Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
          NY. The library is named RPCL-98 and was constructed by partial
          EcoRI digestion of Drosophila DNA provided by the BDGP from the
          isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
          P1 and Est libraries. A more detailed description of the library

```

```

FEATURES
SOURCE
    1..1101
    /organism="Drosophila melanogaster"
    /db_xref="taxon:7227"
    /clone_id="RPCL-98"
    /clone="BACR31A09"
    /note="end : TET3"

BASE COUNT 228 a 157 c 222 g 176 t 318 others
ORIGIN

Query Match 4.5%; Score 56.8; DB 121; Length 1101;
Best Local Similarity 24.5%; Pred. No. 0.0074;
Matches 114; Conservative 133; Mismatches 218; Indels 0; Gaps 0;

QY 1 AANCACCAATCCMATTTGGAGGNAATCCAAATCAATNCCCGANNCCCAATCCAGTTAA 60
    : : : || : : : || : : || : || : || : || : || : || : || : ||
Db 1005 VRMCCAADKDCBDCCARAAMAGDADNAANADGNCMAVAAMCKDKDBKKKKKKKK 946
QY 61 TTAAGTCCAGGTTTGGACATTACCAAAATATGATTCGATGAGTCAATGCAANCG 120
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 945 DKDKMAHBKKKKTKAKTAATAKTGKKKKKKKKTKAAMAKDHANRAAATAAANAAD 886
QY 121 GGTGAATAAAGTACGAATAGTGNAAATGATAGTAGTACAGTGGCATACAGTNT 180
    : : : || || : : || : : || : : || : : || : : || : : || : : ||
Db 885 RKKKARPAATAADADADMDGMAAGRGRTWTMDRMAAGATHTGDTHTTTTH 826
QY 181 CTCNATTTATGATGCTTTTGAATCGAATGAGCAACACAGCGAGTTACATCAATAT 240
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 825 YCCSKYKTTTMAHVCCKTTTGGTAAHTARGGBCWTTWTHTWACAGYAARKAST 766
QY 241 TTGTACACAGAAATAAAGAGACGCTATAGAGACCAATTTGAAGTTGATTTA 300
    : : : || || : : || : : || : : || : : || : : || : : || : : ||
Db 765 TKMAAATAAHHMAAAMMAAATAAARDTCCCAACMAAAMKGGGCAACMAAD 706
QY 301 ATGCTGTAGTAAAGATCATATGATGATGCTTATAGTACTACATTTATTAAT 360
    : : : || || : : || : : || : : || : : || : : || : : || : : ||
Db 705 AHCTTDWATMAVAARAAARAAAMMAAAMWMTDRNMAAMWMTWTTTTDD 646
QY 361 AAAACGTCATTAATGTTTGTAGAAAACATATAGTATCATTTTAAAGTAGTACA 420
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 645 TTTTAMACMAAMWMAAMMAAAMMAAAGAAARAGRGGTTTTTTTTRKATAMC 586
QY 421 TACTACGTACTCAAAATATCTATAACAATTTCATATATTAATCT 465
    : : : || : : || : : || : : || : : || : : || : : || : : ||
Db 585 YTTTATTGMAMARYTBTMAAAMAHCCCAHTTWMAAATTW 541

RESULT 9
AM683426 766 bp mRNA EST 15-JUN-2000
LOCUS NF01IG12LF1099 Developing Leaf Medicago truncatula cDNA clone
DEFINITION NF01IG12LF 5', mRNA sequence.
ACCESSION AM683426
VERSION AM683426.1 GI:7558157
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
          Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
          Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
          Fabales; Fabaceae; Papilionoideae; Medicago.
REFERENCE 1 (bases 1 to 766)
AUTHORS Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
          Flores,H.R., Iman,J.T., Weller,J.W. and May,G.D.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
JOURNAL Medicago truncatula leaf library
COMMENT Unpublished (2000)
          Contact: May GD
          Plant Biology Division
          The Samuel Roberts Noble Foundation

```


Db	491	WATAAWTTTWTWTWAMATTMMKKKKKKKAMADTGAGARATKTDKKKKKGGS	440
RESULT	11		
CNS016JY/C			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
BASE COUNT			
ORIGIN			

[illegible]

DB	681	NENNAVNNNTNNNTNNVTNNVNNTNNTNANNNNNNNNNNNVVNNNTTTTNNNT	622
Oy	669	CCTTGCACTTCATCATGATGCAGACATTTTTGGTTATAGACGATCATCATTGAC	726
		: : : : : : :	
Db	621	TNNNNNNNNVVTTTTTNNAANTNNNVVNTNNNTNTNNVAVVNNNNNNNVV	564
RESULT	12		
LOCUS	CNS00396		
DEFINITION	CNS00396	GSS	03-JUN-1999
	Drosophila melanogaster genome survey sequence Tetr3 end of BAC #		
	BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit		
	fly), genomic survey sequence.		
ACCESSION	AL063921		
VERSION	AL063921.1	GI:4941778	
KEYWORDS	GSS.		
SOURCE	fruit fly.		
ORGANISM	Drosophila melanogaster		
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;		
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 1101)		
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ; BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammosser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2: cn bw sp. the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.		
COMMENT			

FEATURES		Location/Qualifiers
source		1..1101
		/organism="Drosophila melanogaster"
		/db_xref="taxon:7227"
		/clone_1lb="PC1-98"
		/clone="BACROB10"
		/note="end : TET3"
BASE COUNT	201 a 64 c 131 g 202 t 503 others	
ORIGIN		
Query Match	4.3%; Score 54.6; DB 121; Length 1101;	
Best Local Similarity	19.1%; Pred. No. 0.023;	
Matches 105; Conservative 207; Mismatches 237; Indels 0; Gaps 0;		
OY	221 GCGCGATTCATCATCAATTAATTGTTGTAAGAAAGATAAAGAGACCAACCTATATGAGAC	280
DB	392 GGGMATATAMWMMWMTTTTTTTTAAAMAMAMAAATATWMAAMWMAAAAAATTWMAAA	451
OY	281 GAATGATGAGTTGATTTAATGTCGTGTAGTAAATCATCATCATGAGATGCTTAGT	340
DB	452 WMAAMWMTAMWTTTATWMTWMAAAAAAAAATTTTTTTTTTATWTTTATWTTTATWTTT	511
OY	341 TACTCAGATTATTAATAATAAACCGTCATTAAATTGTTTTTTTAGAAAACATATAGTAT	400
DB	512 TTAAMWMAAAAAAAAAMAAAMWATAADTWTWTTTWTWTTTAAWATAAAMCAAW	571
OY	401 CATTTAAATAGTAGTGCATACATACGTAAGTACCAATATCATATCAATTTTCATATATA	460
DB	572 YATTTTYYHYHYTWTWTMTWHTMYHAANHTTWYHYTAAMWMTWHTWMTWMAW	631

0y	461	TTCTTGCGGCGAGGAGAAATTTGSSACGGGAGTAAATGAGCGTGGAGCTGAT	520
Db	632	WTGTTAAUUUUUUUUCMUUUUUNNNNNNAANNAAMTTGTTNNTAANATUUUUUUSAM	691
0y	521	ATGTTCTGATTAATGAGCTGATGAGAGAGATTTGTAAGCGSAGTAAAGTCGGATG	580
Db	692	CMCTNHTGNCUUUUUNNTANTGTTNNNNKANTUUMUUMATUUUMCTGUTUUNNNNNUN	751
0y	581	GGAGAAAGAGATGTTAATGATGAGCAAGATGATGAGCTGATTTGTAATAATGTCASAA	640
Db	752	AUNTTUWMAADAMMMNNNNNANRAAADAAMATGTTNNUNTTUUNNTUUMUUYUUTGUSMC	811
0y	641	ATAGGCGTATGTTAAAGCATGAATTTTCTGCTGCTTGCATGATTCATGATGTCGAGATTT	700
Db	812	TUHCNMUUNHTAUUTGCTUHTNNMMMTWTGTMUUNHTNNNTNNNAMNNNTTTCMWWWNAATTT	871
0y	701	TTTGTTTATAGAGTGCATGATTTGAGTGCATTTATGAGTATATGAGTATGCAACTTSCAAAT	760
Db	872	WATGCMACMTNNNNNNNNNNNNNNAACNANNTNNCMGNNNNCTGNNNTUUNNTGNCMMW	931
0y	761	ATGGTASAA 769	
Db	932	NNNNNNNNNN 940	

RESULT	13
CNS0182P	
LOCUS	
DEFINITION	
	CNS0182P 1101 bp DNA GSS 26-JUL-1999
	Drosophila melanogaster genome survey sequence SP6 end of BAC
	BACN37D10 of DrosBAC library from Drosophila melanogaster (fruit
	fly), genomic survey sequence.

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE
AL108811	AL108811.1	GI:5629115	fruit fly	Drosophila melanogaster	
				Euarthropoda; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroideae; Drosophilidae; Drosophila.	
				(bases 1 to 1101)	

REFERENCE AUTHORS TITLE JOURNAL	COMMENT
1 (bases 1 to 1101) Genoscope. Submitted Submission Directed (22-jul-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)	Web : www.genoscope.cns.fr Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila genome Project (EDGP) - http://www.edgp.ebi.ac.uk . This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billard at CEPH (Centre d'etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.

FEATURES	SOURCE	LOCATION/QUALIFIERS
		1. 1101
		/organism="Drosophila melanogaster"
		/plasmid="pbeloBAC11"
		/db_xref="taxon:7227"
		/clone_11b="DrosBAC"
		/clone="BACN37D10"
		/note="end : SP6"
BASE COUNT	274 a	268 c 128 g 73 t 358 others
ORIGIN		

Query Match	4.3%	Score	54.6	DB	121	Length	1101
Best Local Similarity	21.7%	Pred	No. 0.023				
Matches	126	Conservative	209	Mismatches	243	Indels	2
						Gaps	2
OY	137	AAATGTTGTGNNAAATGATAGTGTAGAACCAAGTGGCCATACAGTTCCTCCNATTTATATATGTC	196				
Db	533	AAAAAAAAAAAAAAAAAGGCGACARACAGCGGRCRGVAGGCGATATATAAAAAAAAAAADD	582				

Oy	197	TTTGATTCGATAGCAACACGCGAGTTTCATCATATTTTGTACGAAAGT -	255
		1 : : : : : : : : : : : : : : : : :	
Db	563	TKDWMDRGAAARRAAAADAAARMAAKWTMAAAAAAAAAAAAAAAAAAGTTKTA	642
Oy	256	AATMAAGAGAACGCTCTATAGACGCAATGAAGCTTTGATTTATGTCTTAGTAAG	315
		1 : : : : : : : : : : : : : : :	
Db	643	AAAAAAGCGMCKTKGCAALDGDGAAADAMWRGRGRRAAARRAAAAADAKRRAAAAA	702
Oy	316	AATCATATCATAGATGCGCTATAGTACTCAGATTATATTAATTAAACGCTCATTAAT	375
		1 : : : : : : : : : : : : : :	
Db	703	AAAAAAAAAAAAADAGGRKKKKDKDGKATTAAMAGRKDMGTAATMMWTDMADT	762
Oy	376	TGTTTTTTAGAAAAACATATAGATATCATTTTAATAGTTAGTACACTACTACTCAA	435
Db	763	WKAATLTDIDKAAAPAAAGRRKRRAARTABDGGRAATRRRAAAGKRRARACARRAARRAA	822
		1 : : : : : : : : : : : : : : : : : : : : : : : : :	
Oy	436	TAACTCTAACAAATTTTCATATATATCTTCGCGGAGCGTGAATCCCAACGCGAG	495
		1 : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Db	823	DDRDMDMAAAAAAAAAAAATTTWRDRKMDMDMDTTRMDDDTTAAMWDARAARRRRR	882
Oy	496	TAAATTAAGCCTGCGACCTCTAATATGTTTCATTTAGTGCCTGATCTAGTAGATTC	555
		1 : : : : : : : : : : : : : : : : : : : : : : :	
Db	883	RRRRRRRAARAARAADDTDKRMDATTDKDTTWTWTTDDDDMKAKRD - RMAAARAD	941
Oy	556	AGAGCCGACAGTTAAAGCTGCGATGGGAGACAAGATGTTAATTTCGACAAAGATAAGT	615
		1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Db	942	GAMWRPBARADDMDAKDDDKDKWKGGRGKRGDKKRMWDKGTGKKDDDDMDKWT	1001
Oy	616	AGCGTATTTGTAATAAATGTACAAATAGCGTTATTTAACGATAATTTTTCCTTC	675
		1 : : : : : : : : : : : : : : : :	
Db	1002	DMMMWMTRTKMDWMDGDRGRGRTRRKMWAMWRAAMARADDTDGKDTPTADKDRKTD	1061
Oy	676	ATCTTAATTCGATGAGAGATTTTGTGTTTAAAGG	715
		1 : : : : : : : : : : : : : : : : : : : : : : : : :	
Db	1062	TKRDGDDMRKDRKDKRRDKGDKTKKADATPMDDDARPM	1101

RESULT	LOCUS	DEFINITION
14	CNS008HI	961 bp DNA
	CNS008HI	GSS
	CNS008HI	03-JUN-1999
	CNS008HI	Drosophila melanogaster genome survey sequence TENG end of BAC #
	CNS008HI	BACR17001 of Rpcc-98 library from Drosophila melanogaster (fruit
	CNS008HI	fly), genomic survey sequence.

ACCESSION	AL051882
VERSION	AL051882.1
KEYWORDS	GI:4933734
SOURCE	GSS.
ORGANISM	fruit fly.
	Drosophila melanogaster
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	1 (bases 1 to 961)

ORGANISM	REFERENCE	JOURNAL	COMMENT
Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 961) Genoscope. Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr)			Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/TheBDGP/Drosophila melanogaster BAC library was prepared by Kazuhiro Oosagawa and Aatón Mannosser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPC1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the library p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .

Location/Qualifiers

